2005-06

Model process validation : an analysis of performance-based pricing programs

Theorgood, Pamela S.

Monterey, California. Naval Postgraduate School

http://hdl.handle.net/10945/1911

NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

MODEL PROCESS VALIDATION:
AN ANALYSIS OF PERFORMANCE-BASED PRICING PROGRAMS

by

Pamela S. Theorgood

June 2005

Principal Advisor:	Philip Candreva
Associate Advisors:	John E. Mutty
	Keebom Kang

Approved for public release; distribution is unlimited

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			*Form Approved OMB No. 0704-0188*
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington DC 20503.			
1. AGENCY USE ONLY *(Leave blank)*	**2. REPORT DATE** June 2005	**3. REPORT TYPE AND DATES COVERED** Master's Thesis	
4. TITLE AND SUBTITLE: Model Process Validation: An Analysis of Performance-Based Pricing Programs			**5. FUNDING NUMBERS**
6. AUTHOR(S) Theorgood, Pamela S.			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000			**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A			**10. SPONSORING/MONITORING AGENCY REPORT NUMBER**
11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.			
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.			**12b. DISTRIBUTION CODE**

13. ABSTRACT (maximum 200 words)

An objective of the Government Performance and Results Act of 1993 and the President's Management Agenda of 2002 has been to integrate performance measurement with resource allocation (budget) decisions. The Department of the Navy has made the use of performance-based pricing models a central part of their efforts to meet these objectives. Using a comparative case methodology, this thesis examines the three models in the Department of the Navy that have been fully accredited. The formulation of these models, the verification, validation, & accreditation (VV&A) process, and their actual use in resource allocation decision making were analyzed against the standards of Navy guidance and industry best practices. In addition to 15 recommendations to strengthen the VV&A process, the research concludes that problems with integration prevent the department from reaching the goal of performance-based budgeting. Improvements can be made in (1) creating conditions to allow the models to better incorporate the effects of naval transformation, (2) assuring knowledge about these models transfers across organization and over time, and (3) linking the budget justification material with the justification used for programming decisions.

14. SUBJECT TERMS Performance-based Pricing Models, Verification, Validation, and Accreditation (VV&A), Performance-based budgeting, Planning, Programming, Budgeting and Execution (PPBE), Modeling and Simulation (M&S), Performance Measurement			**15. NUMBER OF PAGES** 137
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	**18. SECURITY CLASSIFICATION OF THIS PAGE** Unclassified	**19. SECURITY CLASSIFICATION OF ABSTRACT** Unclassified	**20. LIMITATION OF ABSTRACT** UL

NSN 7540-01-280-5500 Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release; distribution is unlimited

MODEL PROCESS VALIDATION: AN ANALYSIS OF PERFORMANCE-BASED PRICING PROGRAMS

Pamela S. Theorgood
Lieutenant Commander, Supply Corps, United States Navy
B.S., Indiana University of Pennsylvania, 1993

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF BUSINESS ADMINISTRATION

from the

NAVAL POSTGRADUATE SCHOOL
June 2005

Author: Pamela S. Theorgood

Approved by: Philip J. Candreva
Principal Advisor

John Mutty
Associate Advisor

Keebom Kang
Associate Advisor

Douglas A. Brooks
Dean, Graduate School of Business and Public Policy

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

An objective of the Government Performance and Results Act of 1993 and the President's Management Agenda of 2002 has been to integrate performance measurement with resource allocation (budget) decisions. The Department of the Navy has made the use of performance-based pricing models a central part of its efforts to meet these objectives. Using a comparative case methodology, this thesis examines the three models in the Department of the Navy that have been fully accredited. The formulation of these models, the verification, validation & accreditation (VV&A) process, and their actual use in resource allocation decision making were analyzed against the standards of Navy guidance and industry best practices. In addition to 15 recommendations to strengthen the VV&A process, the research concludes that problems with integration prevent the department from reaching the goal of performance-based budgeting. Improvements can be made in (1) creating conditions to allow the models to better incorporate the effects of naval transformation, (2) assuring knowledge about these models transfers across the organization and over time, and (3) linking the budget justification material with the justification used for programming decisions.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

First, I would like to thank the Naval Postgraduate School Conrad Scholar Committee, Vice Admiral (Ret) Thomas Hughes, Captain (Ret) John Mutty, and Professor Joseph San Miguel for giving me the opportunity and support to perform this research.

Next, I would like to thank the wonderful team in OPNAV N81 for their assistance in gathering the data for this research. Specifically, the assistance and advice from Mr. Andrew Gallotta, LCDR Ellen Evanoff, and CDR Jane Conley (N4) were an invaluable contribution to this research. As well, a special thanks goes to the various individuals interviewed for the thesis.

Professors John Shank, Nick Dew, Mary Malina and the remaining faculty in the Graduate School of Business and Public Policy deserve special recognition for giving me the vital tools to approach this thesis.

And most of all, I would like to thank my advisors, CDR Philip Candreva, Captain (Ret) John Mutty, and Dr. Keebom Kang for their expertise and guidance. Your commitment to excellence and unmitigating willingness to use up all the red ink on the Monterey Peninsula has propelled me to reach a higher goal than originally anticipated. I have learned much more from this than how to write a thesis and I am honored to have had the opportunity to work with you.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

A. PERFORMANCE MANAGEMENT IN THE GOVERNMENT

There has been much contention since the 1990's on improving government performance. Reports abound evincing the need for better financial management and closer scrutiny of government agencies in carrying out their fiduciary responsibilities. Federal agencies have been confounded for years by the numerous accounts of erroneous payments and benefits amounting to billions of wasted taxpayer dollars. The United States Government Accountability Office (GAO), the accountability and auditing agency of Congress, has voluminously cited the financial management weaknesses and performance shortfalls in government programs. In 1999, Medicare was estimated to have improperly paid over $13.5 billion in uncovered, unnecessary claims (GAO, 2000). The Intercity Passenger Rail Service (Amtrak) has reported net losses in the amount of $762 million to $845 million between 1994 and 1998 (GAO, 1998). And in the Washington Post, GAO announced the inadequacy of government records to perform an effective evaluation of financial soundness. (Lee, 2004) The Department of Defense (DOD) has also received its share of criticism in the performance and financial management of its programs. DOD has experienced difficulty in determining the results achieved from the expenditure of millions of dollars of appropriated money. In 2002, the Defense Department could not track or account for $1.1 billion of money earmarked by Congress for spare parts (GAO, 2003). New initiatives and programs were often implemented without verifying the existence and effectiveness of similar programs resulting in duplicative efforts and an inefficient use of resources (OMB, 2002). And since the Chief Financial Officers Act of 1990 was introduced, the DOD has yet to receive a clean, unqualified audit opinion.

The Government Performance and Results Act (GPRA), enacted in 1993, attempts to resolve these issues and pave the way for more effective resource management practices in federal agencies. The Act requires agencies to develop and use performance measures and goals as a means to achieve the objectives prescribed in the program's annual plan. The annual plan specifies the targeted level of performance required to support the organization's five year strategic plan. GPRA acts as a means to hold the agencies accountable for the results of their programs. The output and outcome performance measures create the opportunity for program managers to evaluate whether their programs are attaining the desired objectives. These performance indicators properly aligned with program activities and full costing data produce the ability for managers to determine the appropriate level of budgetary resources necessary to achieve program objectives. Hence, GRPA articulates the need and essential characteristics of performance-based budgeting. Dollars and program performance are linked to bring about a controlled improvement in overall program management.

The President's Management Agenda (PMA) for fiscal year (FY) 2002 continues in the tradition of the Government Performance and Results Act and implements the Budget and Performance Integration (BPI) Initiative. The BPI, one of the five government-wide program initiatives to reform financial management performance in the Executive branch, describes the need for better information and better control over resources using a performance-based budgeting framework. The initiative directs government agencies to focus on the results of their programs instead of the process (OMB, 2002). Budgetary resources are to be allocated to programs based on proven performance and achievement of targeted results.

Management Initiative Decisions (MID) 901, 910, and 913 represent DOD's response to the President's Management Agenda and the requirements of GPRA. They direct the course the Defense Department will take to comply. The MIDs provide the tools for transforming the DOD into a more accountable and effective organization capable of meeting the needs of the future with higher quality performance intelligence. The Initiatives establish a balanced scorecard for risk management, describe the reporting and assessment of performance metrics, and implement the two year Planning, Programming, Budgeting, and Execution (PPBE) process. The balanced scorecard communicates the strategic aim of the Defense Department, allowing resource allocation decisions to be strategically linked with program and budget development (DOD, 2003a).

The Department of the Navy (DON) links performance with resources by employing performance-based pricing models as one of its primary tools to comply with the President's Budget and Performance Integration Initiative and the various Management Initiative Decisions (DON, 2002c). "Predictable and accountable relationships between resource levels and outcomes" (DON, 2002c) using proven business practices outline essential attributes of the models. Performance pricing models assist the Department in making resource allocation decisions, showing the impact from budget increments or decrements, and supporting budget development. However, before these models are authorized for official use, they must successfully satisfy the criteria in the verification, validation, and accreditation (VV&A) process.

B. OBJECTIVES

If one accepts the premise that linking resource allocation decisions to measures of performance will lead to better decision-making, then the recommendations in this study should lead to improvements in the Department of the Navy. The objective of this research is to improve the management of the department by showing how performance-based pricing models could be better used for program and budget decision-making. The models currently in use will be described in terms of how they are used and how they were built compared to the policy guidance. Further, relying on the academic literature, recommendations are made for how DON might strengthen the VV&A process and create conditions to allow the models to evolve within the Navy's transformation initiatives.

C. RESEARCH QUESTIONS

The primary research question answered by this study is: How can the Navy improve the VV&A process and the integration of performance-based pricing models in program and budget development cycles?

To reach this objective, the following supporting research questions were explored: (1) What is the status of all program performance models that have entered the validation process, (2) How were the fully accredited models built and by whom?, (3) What model costing methodologies were used, (4) How did the verification, validation, and accreditation (VV&A) process work, (5) How are these models used for program and budget development, and (6) What were the key strengths and weaknesses of the three fully accredited performance pricing models?

D. SCOPE OF THESIS

This research assessed the current status of the performance models to include factors such as the number of models that have been developed, the stage of the VV&A process they are in, and what level of accreditation the models have achieved. This study analyzed the use and VV&A process employed for the three fully accredited models in comparison with applicable guidance. The vision of accountability expected from integrating performance and budget information through performance models was examined along with the potential risks to achieving that vision. Numerous recommendations were provided to improve the VV&A process and mitigate the causes of the identified risks.

It was not the intent of this thesis to perform parametric and statistical analysis as to the validity of the models under review. Nor was it the aim of this research to verify, validate, or accredit any model selected for review. This was purely a face value analysis of the methodologies incorporated in the three fully accredited programs modeled for budget and performance integration. The intent was to analyze processes and procedures.

E. METHODOLOGY

This study employed a comparative case study methodology with case data gathered using a theoretical proposition strategy and analyzed using a pattern-matching technique (Yin, 1994). The study set out to determine whether the models were built and used according to the policy guidance. It further examined that policy guidance to see if it could be improved. Then, returning to the models, recommendations were developed for better design and use. While all (n=41) models were concisely evaluated and considered for the study, the three which are fully accredited constituted the comparative cases.

This research involved the examination of government policy and directives on results-oriented, performance-based management and performance-based pricing models using documentation and archival records. The study also involved the evaluation of the verification, validation, and accreditation (VV&A) procedures required for the authorized employment of DON performance-based pricing models in the programming and budget development process. Data to support this evaluation were derived from documentation and interviews with model developers and those who conducted the VV&A process. Further, relevant DON and DOD guidance on VV&A procedures were assessed and evaluated against VV&A procedures found in the business modeling community. The research included (1) a comparison of the roles of the three fully accredited performance models for program and budget development with prescribed guidance; (2) an analysis of the verification, validation, and accreditation of these three Navy performance pricing models; (3) the identification of risks to achieving the GPRA and PMA vision of performance-based budgeting through modeling and simulation; and (4) the provision of recommendations for the improvement of the development and utilization of performance-based pricing models, the VV&A process, and the implementation of future performance-based pricing models for programming and budget development.

F. ORGANIZATION OF STUDY

Chapter II of the thesis consists of background material. It begins with a discussion of relevant government policies and directives pertaining to performance management and performance-based budgeting. This reveals how the performance-based models were supposed to be utilized in the programming and budget development process. Next, it looks at the verification, validation, and accreditation guidance initiated in the Defense Department and reviews VV&A literature from the professional modeling community. This brings out how the VV&A process for performance models was supposed to work and identifies any gaps between DOD policy and commercial practices. Chapter III evaluates the actual uses of the models under review and identifies any divergence from established directives and initiatives. In Chapter IV, the VV&A activities applied to the three fully accredited models are analyzed and compared with applicable guidance and standard practices. Chapter V reviews the goal of GPRA and the President's Budget and Performance Integration Initiative and explores the potential impediments to achieving the vision. This chapter also addresses the themes associated with performance-based budgeting and colligates the performance models with the real intention of this form of budgeting. The goal of this chapter is to identify possible risks to goal achievement and provide recommendations to palliate those risks.

II. FOUNDATION FOR ANALYSIS

This chapter summarizes the literature regarding performance measurement and budgeting in the federal government leading to the implementation of performance-based pricing models. Next, the verification, validation, and accreditation processes used in DOD and in the business modeling community are discussed. The chapter ends with a current snapshot of status of the performance-based pricing models developed for program and budget development in the Department of the Navy. The purpose of this chapter is to provide the foundation for the analysis to follow in subsequent chapters.

A. PERFORMANCE-BASED BUDGETING

1. Government Performance and Results Act of 1993

Performance budgeting is a performance management concept with a long history in the public sector. Largely inspired by the sustained success of the performance measurement and management system implemented by the city of Sunnyvale, California, as well as David Osborne and Ted Gaeblers's ideas in *Reinventing Government,* the Government Performance and Results Act (GRPA) of 1993 legislates the use of performance information in the strategic planning of the federal budget (Mercer, 1994). The GPRA mandates the establishment of performance indicators and performance goals to delineate the targeted level of performance a program is being funded to achieve (Congress, 1993). Linking performance targets with resources elicits the capability for program managers to express the expected results to be attained by a program given a certain amount of budgetary dollars. Infusing performance measures and indicators in the budgeting process adds the additional ability of evaluating how effective a program is in achieving the targeted outcomes. By promoting this new focus on results, these outcomes are expected to be linked and aligned with the agency's annual and five year strategic plan also being required by the GPRA (Congress, 1993). Devising the long term strategic roadmap for an agency and cascading it down to the program level via annual plans fosters the increased likelihood that programmed and budgeted resources are being executed in a manner supportive of the overall vision of the organization. Resource allocation decisions need to be aligned with program goals which should be collimated with the intended direction of the agency. The Office of Management and Budget (OMB) states that, "the agency should develop a 'performance budget' organized like its Strategic Plan that matches resources with outputs and justifies resources requested by the effectiveness at influencing the desired outcomes" (Mercer, 2003, p. 5). This represents one facet of the performance-based budgeting concept.

Performance-based budgeting also requires the identification of well-conceptualized and strategically aligned performance measures and metrics on routine processes. The cost of these routine processes undergirds the utility in implementing a performance-based budgeting system. GPRA specifically focuses on the results side of the performance budgeting equation. However, the intention of GPRA is also for federal agencies to implement the "basic building block of a sophisticated performance-based budgeting and management system, the cost per unit of activity" (Mercer, 2001, p. 10). As a critical component of performance management, performance-based budgeting is intended to diagnostically reveal the intricate interplay between resources, fully costed activities, and targeted performance outcomes (Mercer, 2002) In performance budgeting, day to day tasks or activities are funded vice the traditional program line item object class accounts. The budget is structured with performance language, clearly showing the required activities that need to be performed or executed in order to achieve a certain level of output. The outputs achieved determine the overall outcome or result of the program. Therefore, "managers must have accurate and timely cost and performance information to manage their resources most effectively" (Mercer, 2002, p. 2). They must understand the costs of the efforts involved in reaching the targeted results of their programs.

The budget costs for the activities are recommended to be calculated using the managerial cost accounting methodology of activity-based costing (Mercer, 2002). Activity-based costing (ABC) "captures the current and full costs of performing an activity, provides a context for establishing and monitoring performance measures, and provides the link between activity modeling and economic analysis" (Office of Information Technology, 1995). 'Accurate and timely' costing information linked to program performance processes structures a budgeting system that shows how resource allocation decisions would produce strategic results. Linking dollars to tasks, performance budgets are developed around required performance measures and activities that complect with program objectives. This type of transparency introduces the opportunity to analyze the efficiency and cost-effectiveness of government programs (Mercer, 2001). In a pure performance-based budget, performance is not only measured by achieving planned programming levels and outcomes, but also by improving productivity. Achieving program targets is an important goal, but achieving that same targeted outcome at reduced costs represents a much higher level of performance. The government strives to recapitalize on savings generated from efficiently managed programs to meet competing priorities. Another indicator of a well managed, high performance program is the generation of a higher level of output for the same or reduced amount of inputs (costs). Attaining shorter processing time or lower unit costs clearly reflects improved management practices of the total program. Disaggregating the budget into essential program activities and devising strategic performance measures and cost drivers is an extremely information intensive task. However, "without good cost data, comparisons between competing requirements cannot be made adequately, accurate budgets cannot be developed, and allocations will at best be less than optimal (and even may be antithetical to strategic objectives)" (Candreva, 2004, p.12).

In 2001, the General Accounting Office (GAO) developed a performance-based, results-oriented budgeting framework to assist agencies in integrating performance into their budgeting processes (GAO, 2001). The outcome-focused structure enables agencies to make better use of available resources to accomplish agency goals. Program Managers can use this tool to successfully increase and utilize performance intelligence with budgetary resources. As portrayed in Figure 2.1 below, four mutually related themes are associated with effectively managing for results. Theme one is framed around the desire to allocate budgetary resources based on the program's capability of achieving a targeted performance outcome. Aligning performance information with the budgeting process provides program managers the ability to formulate impact statements concerning funding decisions. The practices in theme two premise on the notion that agencies that based their budget estimates on the most up-to-date and reasonable assumptions would be better equipped to make tradeoffs between covering cost increases and meeting other programmatic needs. Theme three asserts that agencies should display, by program activity, the funding being applied to achieve the performance goals and indicators for that activity and account for both direct and indirect costs of its programs and associated goals (GAO, 2001). Causal relationships between resource allocations and program performance become more transparent under this framework. Theme four incorporates the Japanese principle of *kaizen*, or continuous improvement. Program managers perform gap analysis between the results achieved by their program versus the results targeted and strive to better their performance. Effectively and strategically employing these four themes in concert allows for improved government performance through a results-oriented, performance-based budgeting system.

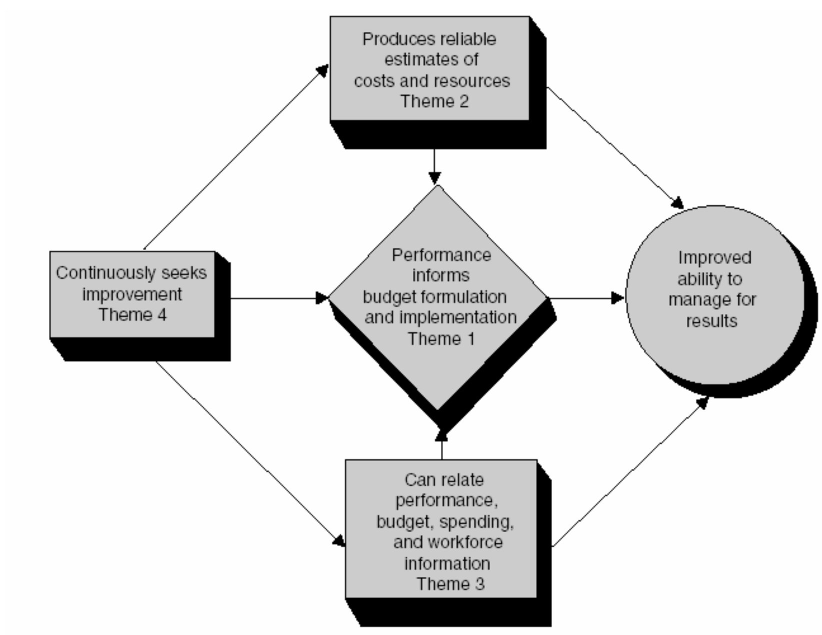

Figure 2.1 Results-Based Budgeting Framework (From: GAO, 2001)

2. President's Management Agenda

Government likes to begin things—to declare grand new programs and causes and national objectives. But good beginnings are not the measure of success. What matters in the end is completion. Performance. Results. Not just making promises, but making good on promises. In my

> Administration, that will be the standard from the farthest regional office of government to the highest office of the land. (OMB, 2002, p. 3).

The President's Management Agenda for fiscal year 2002 expresses the need for improved financial performance in governmental programs through a results-oriented, performance-based budgeting structure. In many aspects, the PMA acts as a tautological representation of the Government Performance and Results Act of 1993. Through the establishment of five government-wide mutually reinforcing initiatives, the President communicates the agenda to reform the management of government programs through better, more goal-oriented business practices. The first management initiative discusses Human Capital Strategy (HCS). Each agency's human capital strategy should be aligned with the overall mission of the organization so that the agency employs the "right people, in the right places, at the right time" (OMB, 2002, p. 14). The second Presidential initiative focuses on Competitive Sourcing which aims to increase the competition in government to improve performance and generate savings from market-based activities (OMB, 2002). Improving Financial Performance with 'accurate and timely financial information' is the main thrust of the third Management Initiative. Fourth, through the Expanded Electronic Government Initiative, the President asserts the need to deliver greater productivity through more citizen-friendly electronic solutions. (OMB, 2002) The success of these four Presidential reformations relies on the government's ability to achieve the fifth management initiative, the Budget and Performance Integration Initiative (BPI).

The BPI designs to link resource allocation decisions with proven program effectiveness through the implementation of performance-based budgets. Program managers have to provide evidence that their programs are actually accomplishing their goals and achieving desired results. Better performing programs are expected to receive a better share of the scarce budgetary resources. The BPI defines better performance as "an assessment of the expected outcome relative to what is actually being achieved" (OMB, 2002, p. 30). In essence, performance (outcome) measures were to be established so that agencies could determine the effectiveness of their programs and be fiscally rewarded for achieving the targeted level of performance. Underperforming programs that persistently failed to achieve the expected outcome measure are to be identified with an eye to cutting their funding, redesigning them, or eliminating them altogether (OMB, 2002).

The BPI also acknowledges the need for more accurate cost measurement systems so that agencies can gain better performance intelligence on the results of their programs and ascertain methods of improvement. Linking budgetary costs and resources with program performance evokes the possibility of improving the performance management of the various agencies and the government as a whole. Yet, the crux of this budget and performance integration idea and the management initiatives of the President's Agenda emphasize transforming the business of government from a process-driven entity to a results-oriented, outcome-focused enterprise.

3. Defense Department Response

Management Initiative Decisions (MID) 901, 910, and 913 represent the guidance implemented by the Defense Department to comply with the requirements of the PMA and GPRA. They are aimed at providing the tools for transforming the Department of Defense into a more accountable and effective organization capable of meeting the needs of the future with high quality performance information. Collectively, these Initiatives focus heavily on the development and linkage of performance measures, goals, and organizational strategy.

MID 901, "Establishing Performance Outcomes and Tracking Performance Results for the Department of Defense," establishes a balanced scorecard and a risk management framework as a way of aligning near and intermediate program results with long term strategic goals (DOD, 2002a). It is a management tool that enables managers at all levels to monitor the results of their programs. "Every measure selected for a balanced scorecard should be an element of a chain of cause-and-effect relationships that communicates the meaning of the business unit's strategy to the organization" (Kaplan and Norton, 1996, p. 149). Defense agencies are expected to ensure their program goals and performance metrics are congruent with the outcome measures delineated in the Department scorecard and overall strategy. Balanced scorecards enable senior managers to assess whether their programs are creating value for current and future capabilities

(Simmons, 2000). In DOD, the executive scorecard communicates the corporate strategy throughout the department, framed around four risk management quadrants; force management risk, operational risk, institutional risk, and future challenges risk (DOD, 2002a). This is a Defense adaptation of Kaplan and Norton's balanced scorecard comprising the four quadrants relating to Financial Perspective, Customer Perspective, Internal Business Perspective, and Innovation and Learning Perspective (Simmons 2000).

The Force Management Risk Quadrant measures the sustainability and quality of the forces available in the Defense Department. The Operational Risk Quadrant tracks the combat readiness of current forces and their capability to respond to external threats. Institution Risk addresses the financial acquisition management and logistical aspects of DOD. It also measures the organization's ability to manage direct and indirect costs. The Future Challenges Risk Quadrant evaluates the innovative and transformational results of Defense programs. The Decision also designates the Under Secretary of Defense for Personnel and Readiness (USD P&R) as the lead for complying with the President's Management Agenda (DOD, 2002a). As such, he establishes supporting scorecards for each of the five Management Reforms initiated in the PMA. The balanced scorecards assist program managers in designing strategic performance measures and metrics for programming and budgeting decisions by providing the strategic vision the department.

MID 910, "The Budget and Performance Integration Initiative," expands on the BPI guidance in the President's Management Agenda and reinforces the Executive Scorecard established in MID 901. The Defense BPI "instructs the Components to associate performance metrics with resources in the President's Budgets," (DOD , 2002b, p. 1) and describes how the Department of Defense intends to improve its rating on the President's Executive Scorecard. Budget and Performance Integration also directs Defense agencies to incorporate performance metrics into budget justification material to clarify the relationship between program plans, targeted outputs, and budgetary resources. The BPI specifically states that "congressional justification material should indicate how individual programs link to the performance outcomes listed in MID 901 for the DOD Balanced Scorecard (BSC) risk management" (DOD, 2002b, p. 3). It also identifies the requirement to verify and validate all performance metrics incorporated in budget justification documents. The overriding objective of DOD's Budget and Performance Integration Initiative is to align resources with program strategic plans that were linked to the agency's overall strategic focus (DOD, 2002b).

MID 913, "Implementation of a 2-Year Planning, Programming, Budgeting, and Execution Process," initiates a two year budgeting process with increased emphasis on executing the budget plan in the second year (DOD, 2003). The overriding objective is to inform resource allocation decisions through an integrated programming and budgeting process. The transformation to the Planning, Programming, Budgeting, and Execution (PPBE) process is the latest change to the Planning, Programming, and Budgeting System (PPBS) that has guided the defense budgeting process since the 1960's. PPBE provides a framework for strategy development, planning for military capabilities, and making resource allocation decisions based on those targeted capabilities to achieve the Defense mission. During budget review and budget execution, program performance metrics that the Components submit as part of the budget estimate submission are used to inform resource decisions and measure the performance of funded programs. Program managers assess the performance quality of their particular programs based on the analysis of targeted and achieved output information. Necessary adjustments are made during this phase to reach the performance goals established during the programming and budget development process. MID 913 clearly shifted the focus of DOD Components to a results-oriented, performance-driven financial management system.

4. Department of the Navy's Budget and Performance Tool

In the Department of the Navy (DON), the Chief of Naval Operations (CNO) cascades the DOD vision and PMA Initiatives throughout the sea services. In accordance with the Sea Enterprise concept and the Department of the Navy Resource Management Initiatives, the CNO encourages all levels of the organization to "continuously strive to minimize the amount of resources needed to achieve success in current operations" (DON, 2002c). DON financial resources are to be aligned with performance. As part of the resource management concept and as much as possible, performance-based pricing models are to be employed to determine the short term resource requirements of the Department. "Sea Enterprise" lists the desired model characteristics as follows:

1) Integrate budget and performance information in accordance with the President's Management Agenda, by establishing a predictable and accountable relationship between resource levels and outcomes to be achieved.
2) Allow for DON Headquarters performance specification and process oversight, but be actively managed by the lowest organization level responsible for functional management of the program area.
3) Incorporate core efficiency achievement, proven business practices, and savings in accordance with Sea Enterprise expectations, with accommodation of specific initiatives in the near-term and improvement targets/investment benefits in the long term.
4) Avoid the false precision of over-specification of requirements, and promote flexibility of execution within the general parameters of each model.
5) Allow the Department to promote accountability for results, and incentivize managers to be successful at lower than projected costs. (DON, 2002c).

The Department of the Navy has accredited over forty performance-based pricing models to improve efficiencies in program areas and to recapitalize on the savings identified through results-based budgeting practices. These models are designed to enable program managers to determine the capability and performance levels attainable given a certain level of funding and comply with the aforementioned Initiatives. Comptrollers and FMB Analysts should also be able to use these models to "confidently project costs and consequences of proposed budget actions and to support budget development" (Myers, 2004). The following definitions apply to the DON performance-based modeling concept:

a) Model – a physical, mathematical, or otherwise logical representation of a system, entity, phenomenon, or process.

b) Performance model – a model that determines requirements for a specific readiness, manpower, training, or infrastructure program that is based on CNO-approved performance goals. In computational form, this is unit times unit cost.

c) Pricing model – a subcomponent of a requirements generation model that produces the unit cost for a specific readiness, manpower, training, or infrastructure program.

d) Requirements generation model (Performance-based pricing model) – defined methodology for determining requirements in support of readiness, personnel, training, or shore infrastructure programs and pricing the specified requirements. (Department of the Navy, 2003a).

At the time of this research, only three performance-based pricing models have received full accreditation; the Aviation Depot Maintenance Model, the Environmental Restoration Model, and the Nuclear Propulsion Technical Logistics Model. The Aviation Depot Maintenance Model, one of the selected models for this study, is composed of two separate requirements models; the Airframe Depot Readiness Assessment Model (ADRAM) and the Engine Depot Readiness Assessment Model. (EDRAM) The following diagram in Figure 2.2 provides the reader with an example of how a performance-based pricing model may be constructed to predict program requirements. These models must undergo a VV&A process to build user confidence in requirements determinations and to "enhance the value of the performance models as both a programming and budgeting tool" (Myers, 2003).

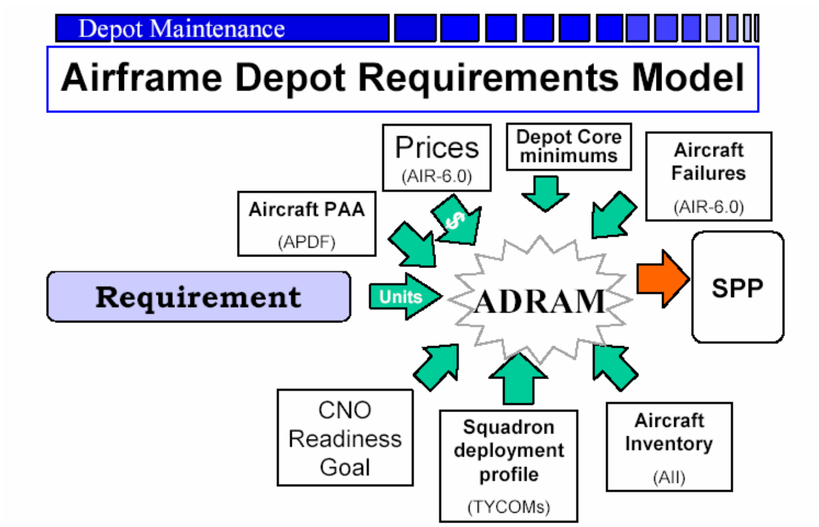

Figure 2.2 ADRAMs Model (From: DON, 2003d)

B. OVERVIEW OF THE VERIFICATION, VALIDATION, AND ACCREDITATION (VV&A) PROCESS

1. Rationale for VV&A

The U.S. Department of Defense is the largest sponsor and user of Modeling and Simulation (M&S) applications in the world (Balci et al, 2000). The DOD has employed models for war gaming and simulation scenarios, major acquisition decisions, human relations and training analyses. In response to the Budget and Performance Initiative, the Sea Enterprise and transformation efforts, and the scarcity of funding resources in the Department's annual budget, performance-based pricing models are being relied on for resource allocation decision-making. With the Department of the Navy's fiduciary responsibility of over $100 billion dollars, it is critical that decision-makers have a high degree of confidence in the ability of these performance models to produce accurate, reliable results. The verification, validation, and accreditation process is designed to establish credibility and confidence in the use of these models for making programmatic decisions and developing a performance-based budget. VV&A is a risk management process engineered to provide an understanding of the capabilities and limitations of the model as well as to mitigate the cost, schedule, and performance risks (Department of the Navy, 2004a). The process builds assurance and trust in the performance-based models' capability of portraying the mission impact of their decisions.

2. VV&A Taxonomy

Confidence in a particular model must be justified before its results are used to make decisions involving large sums of money or risk to human life (Department of the Navy, 1999). All performance-based pricing models must go through a stringent verification, validation, and accreditation process to provide a certain level of confidence to resource decision-makers on the impact of their funding choices. Performance models are primarily used to plan and budget for targeted capability and readiness levels for Navy programs to support the overall strategic plan. They are relied upon to ascertain the most accurate costs associated with the desired capability levels and the impact on readiness due to budget increments and decrements. Only selected program models that have received accreditation are authorized for official use.

1) Verification: is the process of determining that a model or simulation implementation accurately represents the developer's conceptual description and specifications.

2) Validation: is the process of determining the degree to which a model or simulation is an accurate representation of the real world from the perspective of the intended use.

3) Accreditation: is an official determination that a model or simulation is acceptable to use for a specific purpose (Department of the Navy, 1999).

Verification, validation, and accreditation are three separate but interrelated processes interlinked as a system to ensure the results derived from the performance model meet the level of accuracy necessary for its intended purpose.

3. VV&A Techniques

There are no algorithmic, optimized sets of verification, validation, and accreditation procedures or sets of activities that would universally apply to all performance pricing models. However, a study has been done to compile information from a wide variety of sources, including DOD directives and instructions related to M&S management and VV&A (Chew and Sullivan, 2000). The results of this study delivered a set of critical VV&A tasks that should be performed during each phase of the modeling development lifecycle. The effectiveness of the performance model is in direct relation to the effectiveness of the verification, validation, and accreditation process. In aggregate, the professional literature articulates the necessity of drafting a thorough accreditation and V&V plan, beginning the V&V process as early as possible in the development of the required model, and ensuring adequate documentation and coordination is maintained throughout the entire process. As depicted in Figure 2.3, the VV&A team is actively engaged in each phase of model creation and appropriately aligns and performs essential V&V activities prior to the commencement of the next major development activity.

All of these tasks do not apply to every modeling effort, however. The V&V process for each model should be tailored for the intended use, model complexity, size, and acceptability requirements of that particular model. Each individual VV&A effort must strike a balance between cost-effectiveness, responsiveness, and sufficiency to succeed (Department of the Navy, 1999). The direct correlation between model confidence (user trust) and VV&A costs contributes to the selection of V&V techniques to be employed for each model. VV&A tasks should be tailored to high risk performance areas of the model. The only programs encouraged to incorporate all of these confidence building activities are the models being designed to support large dollar or safety of life programs.

The professional literature details the VV&A actions displayed in Figure 2.4 while the DON Modeling and Simulation (M&S) Verification, Validation, and Accreditation Handbook specifies the M&S and VV&A interplay depicted in the diagram below (Figure 2.3).

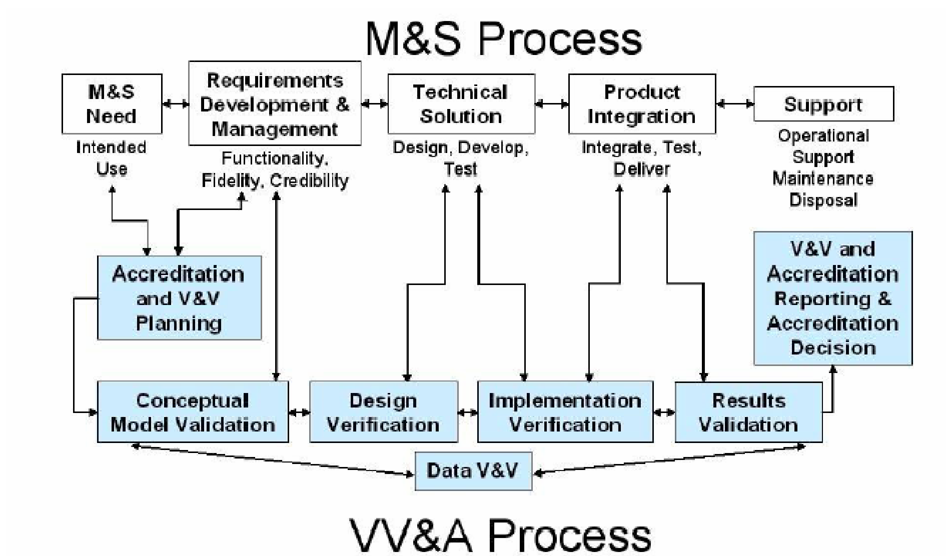

Figure 2.3 DOD Modeling and VV&A Process (From: DON, 2004b)

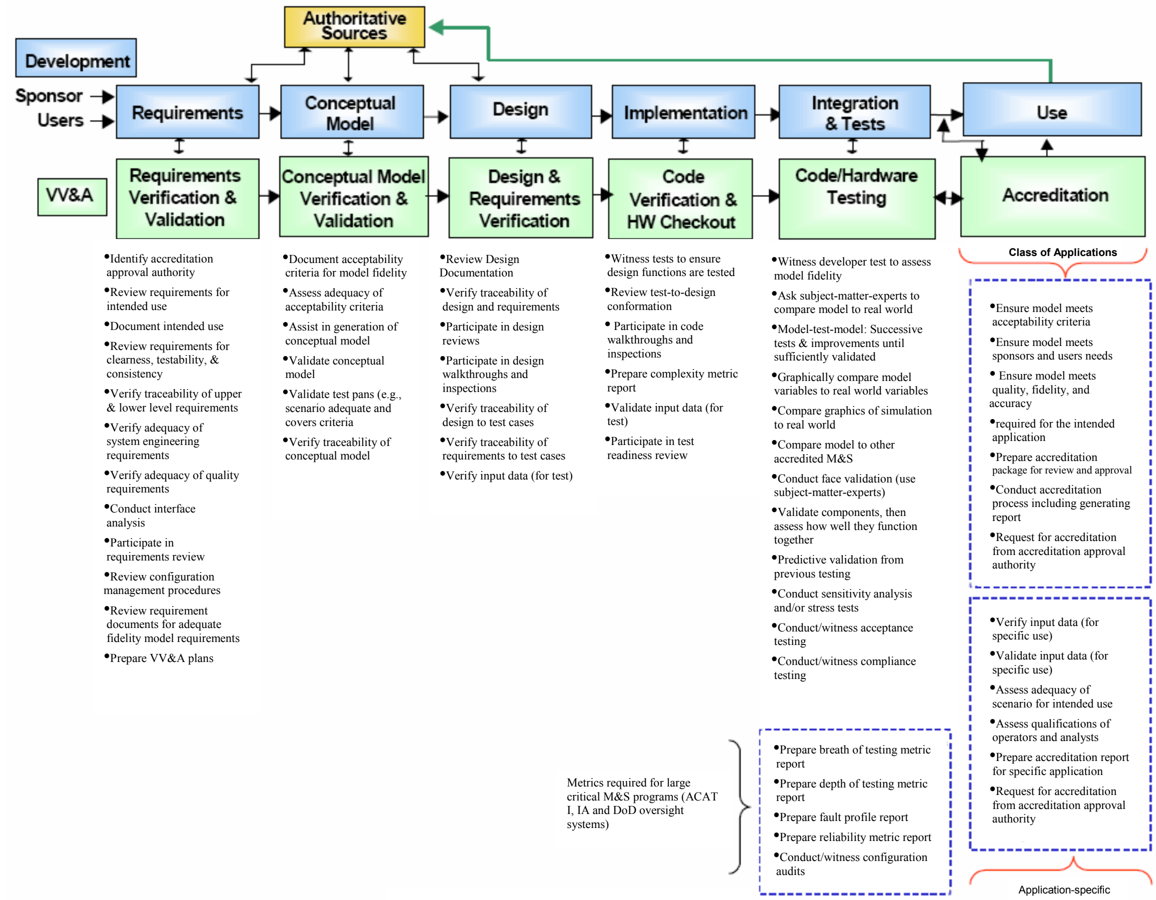

Figure 2.4 VV&A in the Life Cycle of M&S (From: Chew & Sullivan, 2000)

Regardless of which structure the proponents intend to follow, they both will entail many of the same activities and produce the same products.

(1) Requirements V&V: Planning for VV&A is just as essential as implementing the specified techniques. A properly prepared accreditation and V&V plan assists in communicating a clear, well-understood, and complete requirements document. This phase verifies and validates the intended use, the exit criteria for accreditation, and the level of fidelity and detail required for acceptability.

(2) Conceptual Model V&V: This is the creation of the theoretical framework or design of the model based on the results of the requirements V&V. Component parts of the model are analyzed to determine how they relate and interact with each other. The

determination of equations and underlying assumptions are described in this process. Any findings at this stage are reported and resolved before the design of the model can begin.

(3) Design Verification: A detailed design supporting the conceptual model and linking directly to the model requirements is developed. Traceability to the first two phases of model development is key at this stage of the process.

(4) Code Verification: The code and the design of the model must be verified for accuracy and functionality. Input data are continuously validated and any discrepancies need to get resolved prior to full model integration and testing can begin.

(5) Code and Hardware Testing: The code and all the components of the model are integrated and tested for its ability to simulate the real world. There are various levels of testing, covering breadth, depth, and reliability, that vary depending on level of confidence desired in the performance model. There are also various validation techniques to prove fidelity, ranging from face validation to benchmarking techniques. The more reliable the measure selected, the smoother the accreditation process should be.

(6) Accreditation: All V&V and model development documentation is reviewed at this stage of the process and rigorous tests are performed on input data. Data are thoroughly analyzed not only for the accuracy, reliability, and availability from data sources but also for how they must be transformed or manipulated to be used in the model itself. The accreditation agent should be independent from the V&V team and ensure the process is application-specific.

C. CURRENT STATUS OF DON PERFORMANCE PRICING MODELS

Though the available literature suggests that the accreditation acceptability criteria for each model should be tailored to that model, the Department of the Navy accredits performance-based pricing models using the standard exit criteria delineated in DON Verification and Validation template in Appendix A. There are eight separate performance criteria that need to be measured, each with a red, yellow, and green rating scale for scoring. The models are accredited based on the scores they receive in the areas of CNO performance goals, costs for performance capability levels, driver accuracy, modeled components, design, configuration management, feedback loop, and user friendliness.

Depending on the score received, a performance model will qualify for one of the five possible accreditation decision options. Appendix B shows them as full accreditation, accreditation with limitations, accreditation with modifications, additional V&V needed, and no accreditation. Out of almost 50 performance-based models developed for programming and budget development cycles, only three models have actually reached full accreditation with high confidence. They are the Aviation Depot Maintenance Model, the Nuclear Propulsion Technical Logistics Model, and the Environmental Restoration Model. Approximately 16 performance models achieved accreditation with medium confidence and the rest are still working through the verification and validation process. The status displayed in the Table 2.1 represents the most up to date information delivered by the Office of Chief of Naval Operations, N81 Performance Model Accreditation Point of Contact as of 16 November 2004.

Modeled Program	V V Approx		Accreditation		Notes/Remarks
	Due Date	% Complete	Due Date	% Complete	
Sealift, Prepo & Surge	30-SEP-04	100%	31-Dec-04	0%	Completed V&V
Enlisted Accessions	30-SEP-04	30%	31-Dec-04	0%	V&V stage
SELRES Support Model	30-SEP-04	40%	31-Dec-04	0%	V&V stage
Fleet Systems Engineering	30-SEP-04	0%	31-Dec-04	0%	Requires V&V
Cruise Missile	31-OCT-04	10%	31-JAN-04	0%	Development stage
Flight Training (Production/Pricing)	31-DEC-04	60%	31-JAN-04	0%	V&V stage
Specialized Skill Training (Initial Skills)	31-DEC-04	10%	31-JAN-04	0%	V&V stage
Ordnance Support/RSSI	30-SEP-04	60%	31-Dec-04	0%	Funding constraints
Fleet Change Improvement Plan (FCIP)	N/A	0%	N/A	0%	Revert back to LOE
Fleet Modernization Plan (FMP)	N/A	0%	N/A	0%	Revert back to LOE
Weapons Maintenance – Special Weapons	31-AUG-04	100%	30-NOV-04	0%	Accreditation in Progress
PCS	15-JUL-04	70%	15-OCT-04	0%	V&V stage
TACAMO	1-JUN-04	0%	1-SEP-04	0%	Part of FHP program
AMCM Life Cycle Support	30-APR-04	0%	31-JUL-04	0%	V&V required
Navy Nuclear Propulsion	**30-APR-04**	**100%**	**31-JUL-04**	**100%**	**Fully accredited**
Acquisition Management Program	30-APR-04	100%	31-JUL-04	75%	Accreditation in Progress
Operational METOC –T-AGS	30-APR-04	100%	31-JUL-04	100%	Accredited with limitations
Fleet Hospital Program	30-APR-04	100%	31-JUL-04	25%	Accreditation in progress
OMIS	31-MAY-04	100%	31-AUG-04	50%	Accreditation in progress
IA Projection	31-MAY-04	100%	31-AUG-04	75%	Accreditation in progress
Undersea Surveillance (1C3C)	30-JUN-04	75%	30-SEP-04	50%	V&V stage
Submarine Combat Systems (1B2B)	30-JUN-04	80%	30-SEP-04	50%	Accreditation Team requires V&V data
Voluntary Education	30-JUN-04	70%	30-SEP-04	0%	V&V stage
Amphibious Tactical Support – LCAC	15-JUL-04	100%	15-OCT-04	75%	Completed V&V
Air Traffic Control & Combat ID System	31-JUL-04	100%	31-OCT-04	80%	Completed V&V
Weapons Maintenance - Air Launce Missile	31-AUG-04	100%	30-NOV-04	75%	Completed V&V
Weapons Maintenance – Air Launched Ordnance	31-AUG-04	100%	30-NOV-04	75%	Completed V&V
Target Maintenance	31-AUG-04	100%	30-NOV-04	75%	Completed V&V
Weapons Depot Maintenance – Phase II	30-SEP-04	100%	31-Dec-04	75%	Completed V&V
Other Personnel Support PEINSURV	30-SEP-04	100%	31-Dec-04	0%	Completed V&V
Equipment Maintenance – SE Depot Rework	30-SEP-04	90%	31-Dec-04	0%	V&V stage
Air System Support	30-SEP-04	40%	31-Dec-04	0%	V&V stage
Flying Hour Program	29-FEB-04	100%	TBD	100%	Accredited with Limitations
Ship Ops	3-OCT-03	100%	1-DEC-03	100%	Accredited with limitations
Ship Maintenance	24-OCT-03	100%	15-JAN-04	100%	Accredited with limitations
Ship Maintenance Support	24-OCT-03	100%	15-DEC-03	100%	Accredited with

					limitations
Aircraft Depot Maintenance	**19-SEP-03**	**100%**	**15-NOV-03**	**100%**	**Fully accredited**
Engines Depot Maintenance	**20-SEP-03**	**100%**	**16-NOV-03**	**100%**	**Fully accredited**
Depot Maintenance – Weapons	30-DEC-03	100%	14-MAY-04	100%	Accredited with limitations
Spares	31-AUG-03	100%	31-OCT-03	100%	Accredited with limitations
Base Operating Support	31-AUG-03	100%	30-NOV-03	100%	Accredited with limitations
Facilities Investment	30-SEP-03	100%	15-OCT-03	100%	Accredited with limitations
Family Housing	31-JUL-03	100%	30-SEP-03	100%	Accredited with modifications
Naval Construction Force	31-OCT-03	100%	15-DEC-03	100%	Accredited with limitations
Environmental Readiness	N/A	100%	15-JAN-04	100%	Accredited with modifications
Environmental Restoration	**N/A**	**100%**	**15-JAN-04**	**100%**	**Fully accredited**
Force Inactivation (Ships Only)	24-OCT-03	100%	15-JAN-04	100%	Accredited with modifications
Naval Aviator Inventory	15-OCT-03	100%	30-DEC-03	100%	Accredited with modifications
MILPERS	27-JUN-03	100%	31-DEC-03	100%	Accredited with modifications

Table 2.1 Status of DON Performance-Based Pricing Models as 16 November 2004

D. SUMMARY

This chapter has introduced the literature regarding performance measurement and budgeting in the federal government that contributed to the implementation of performance-based pricing models. The rationale and process underlying the verification, validation, and accreditation requirement was also discussed. Finally, the current status of the DON performance-based pricing models was provided. Of particular note is the definition of performance-based models and related terminology were supplied. For the remainder of the study, the term performance model will be used interchangeably with the term performance-based pricing model and requirements generation model. Given this foundation, the next chapter will analyze the use and benefits of the Aviation Depot Maintenance Model, the Navy Nuclear Propulsion Technical Logistics Model (NPTL), and the Environmental Restoration Model.

III. CURRENT USE OF PERFORMANCE-BASED PRICING MODELS

The last chapter discussed the rationale behind using performance-based pricing models and the desired model capabilities. This chapter describes the role of performance models in the Department of the Navy's programming and budget development process. Specifically, the chapter will detail the purpose and functional aspects of two elements of the Aviation Depot Maintenance Model (airframes and engines), the Nuclear Propulsion Technical Logistics Model, and the Environmental Restoration Model. The chapter concludes with a summary of capabilities delivered by the three fully accredited models.

A. INTRODUCTION

Recently, there have been many reforms implemented in the Department of Defense that have impacted the program and budget development system. Through resource allocation decisions, resource sponsors are expected to better manage their programs. Performance measures and balanced scorecards are being developed to hold program managers accountable for the execution of their funded programs. As well, program managers are expected to align the targeted objectives and results of their management efforts with organizational strategy. DOD is shifting from a process-oriented program and budget focus to one centered around outcomes. As expressed by the Secretary of Defense, "the Department currently is pursuing transformational business and planning practices such as adaptive planning, a more entrepreneurial, future-oriented capabilities-based resource allocation process, accelerated acquisition cycles built on spiral development, output-based management, and a reformed analytic support agenda" (DOD, 2003b, p. 6). The requirements generation process has changed from a stove-piped bottom-up structure to one where the requirements originate from a joint, top-down approach. Now, it is more critical than ever for the Department of the Navy to be able to defend and justify its program requirements.

Through the Sea Enterprise Board of Directors (BOD) and Resource Management requirements, the DON now demonstrates its Total Obligation Authority (TOA) resource allocation plans via three pillars; performance-based pricing models, future mission capability alternatives, and Level of Effort (LOE) programs. Figure 3.1, provides an example of the capability-based budget development structure and the increased role played by performance models.

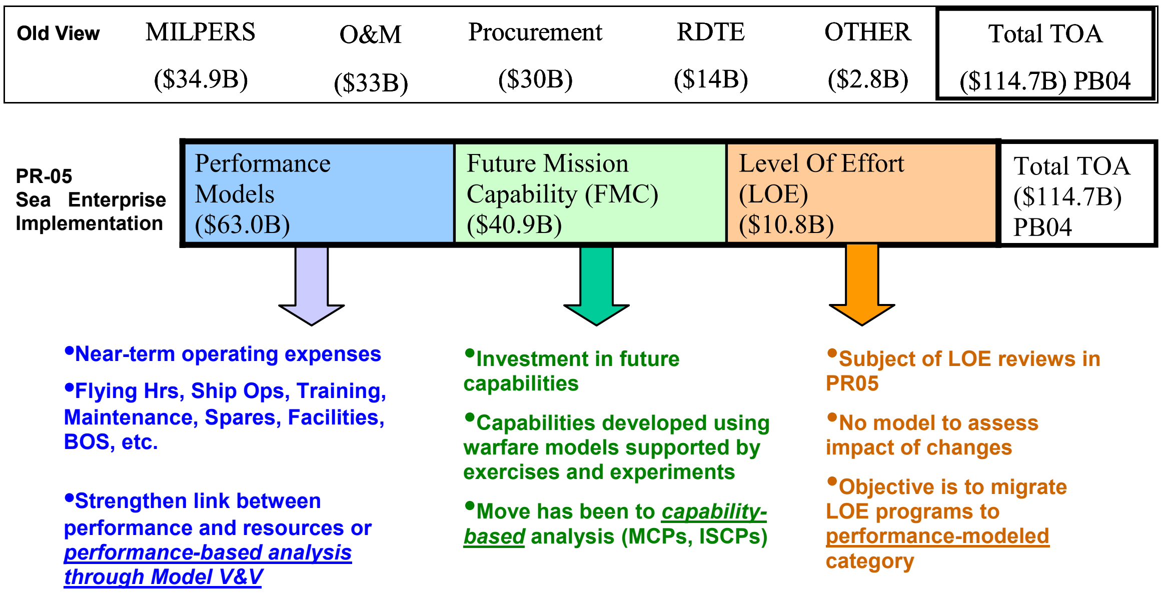

Figure 3.1 DON Capability/ Performance Pillars (From: Sea Enterprise BOD Brief 2003)

Level of Effort programs are those programs that “do not lend themselves to either pricing/performance-based models or capability/investment alternative analysis” (DON, 2003c, p.5). They are primarily based on historical spending data. The future mission capabilities pillar captures the research and development and major investment program accounts and is required to be at least 34% of Navy TOA (DON, 2003). The requirements under this category often derive from campaign analysis modeling and scenarios conducted by resource sponsors. The component comprising the greatest amount of Total Obligation Authority and detailing the critical relationship between cost and performance capability contains the performance-based pricing models.

As stated in Chapter II, the purpose of performance-based pricing models is to “integrate budget and performance information in accordance with the President’s Management Agenda, by establishing a predictable and accountable relationship between resource levels and outcomes to be achieved" (DON, 2002c). The models bring transparency to the linkage between price, program performance, and budgetary resources to achieve CNO goals.

B. INTEGRATION OF PERFORMANCE MODELS IN THE PLANNING, PROGRAMMING, BUDGETING AND EXECUTION (PPBE) PROCESS

In compliance with the Secretary of the Navy's (DON, 2002b) directive to use performance-based pricing models to the maximum extent practicable and to justify program TOA decisions within the three aforementioned program pillars, the Navy is working to integrate performance models in all phases of PPBE. The overarching goal is to develop the most executable, strategically balanced budget through stakeholder consensus and participation. According to the Program Objective Memorandum (POM) guidance for POM-06, there are "six distinct phases of the POM-06 BOD process" (DON, 2004b1, p. 2). They include:

- Sponsor Capability Plans (January-February)
- Integrated Sponsor Capability Plans (February-March)
- Integrated Investment Strategy (April)
- Sponsor Program Proposal (April-May)
- Navy POM Brief (May)
- N82/FMB Budget Formulation (June-August)

The Sponsor Capability Plan (SCP) stage, typically conducted in the first two months of the year, consists of the development of program capabilities by the Resource Sponsors. Per POM-06 guidance, performance-based pricing models assist in the determination of planned readiness and performance levels that each program is capable of achieving based on Military Capability Plans, Total Force analysis, and program reviews. The utilization of accredited performance models for readiness capability planning provides for the ability to present computational, output-based justification and defensibility of planned resource allocation decisions. The Sponsors clearly define the cost driver relationships and assumptions, the key issues associated with various performance levels, and the costs to execute to the desired capability levels. SCPs also address "on-going efforts to control/reduce cost drivers" (DON, 2004b1, p.2) applicable to each program. Once the SCPs are briefed to the 2 and 3 star Board of Directors (BOD) and the CNO's Executive Board (CEB), N8, the Resource and Requirements Assessment Directorate delivers Senior-level feedback to the Resource Sponsors for consideration.

The next phase, the Integrated Sponsor Capability Plans (ISCP), is focused on refining the SCP to incorporate BOD and CEB guidance, as well as fiscal planning information shared between Resource Sponsors. It is during this phase that the various program stakeholders, to include FMB, OPNAV, ASN (RD&A), and the SYSCOMs, interact with Sponsors and program managers to "identify significant pricing issues early in program development" (DON, 2003c, p. 6). Between February and March, FMB coordinates an Investment Pricing Validation Team (IPVT) review to address cost driver issues and cost and pricing increases associated with the program. Additionally, and specifically for performance models, a Model Pricing Validation Team (MPVT) review takes place in the same timeframe to address similar pricing matters consociated with M&S performance plans to ensure that the pricing factors are accurate and defendable. The objective of the ISCP stage is for Resource Sponsors to incorporate this additional information in their previously submitted SCPs and modeled programs to fine-tune their capability goals.

In the third phase, the Integrated Investment Strategy is developed by N80 and N81 OPNAV Directorates to aggregate the various ISCPs and provide fiscally balanced program options to the Resource Sponsors. Upon receiving approval from the CNO's Executive Board, N8 issues the Navy's Program Guidance for Resource Sponsors to adjust their program plans and use performance models to inform trade-off decisions.

Between April and May, fiscally informed Sponsor Program Proposals (SPP) are generated incorporating the program desires of senior leadership, concerns of the approved ISCPs, and the cost and pricing factors resolved during the IPVTs and MPVTs. Performance-based pricing models are used to recalculate program capability levels, based on this information, and output the total costs associated with performing to the readiness plan. The Sponsor Program Proposals are manually input into the Program

Budget Information System (PBIS). PBIS is the integrated program and budget database designed to "consolidate the current programming and budgeting auditing trails (WINPAT and NBTS)… and will be the principle tool through which all stakeholders interact during DON POM/ Budget development" (DON, 2003c, p. 6). Windows Program Analyst's Toolkit (WINPAT) was a separate Navy Future Years Defense Program database that provided programming PPBE phase information for program managers. The Navy Budget Tracking System (NBTS) was the FMB's budget tracking database that did not communicate or interoperate with the WINPAT system. The new combined system, PBIS, allows all stakeholders to view and track the various decisions and plans being made with respect to their programs of interest. After a certain period, the proposal is locked and briefed to the BOD and CEB for assessment.

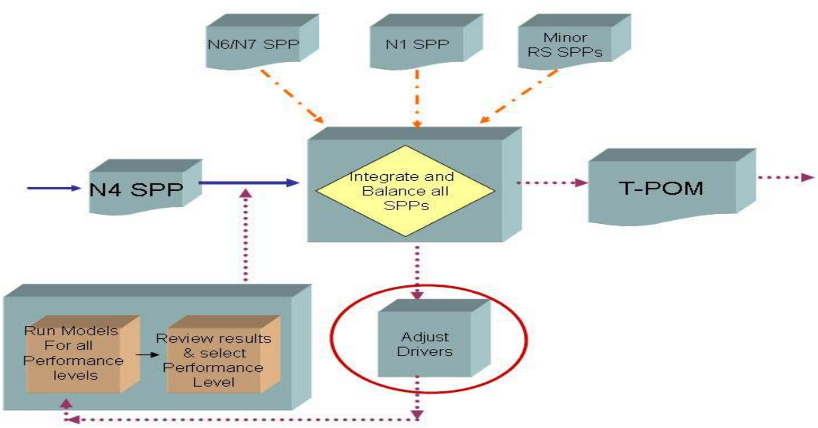

Figure 3.2 Integration of SPPs (From: DON, 2003a)

Based on the updated guidance received from the SPP CEB brief, N80 and N81 build the Navy Program Objective Memorandum. As portrayed in Figure 3.2 above, the Navy integrates the Resource Sponsor SPP's, using the models to readjust performance levels to create a synergistically balanced, strategic Navy programming document. The POM incorporates all of the approved ISCP goals, senior leadership plans, and fiscally informed capability targets for the various programs to execute in accordance with Navy strategy. By the end of May, the POM is fully input into PBIS for FMB review.

Mid-June, emphasis shifts from planning and programming to budget formulation. During budget development, "FMB creates the DON budget after conducting a review including program re-pricing, execution analysis, and re-balancing" (DON, 2003a, p.3). Per POM-06 guidance, performance models could be useful in evaluating the impact on readiness caused by FMB re-pricing and re-balancing before submitting the DON Budget Estimate to the Office of the Secretary of Defense.

C. AIRFRAME DEPOT READINESS ASSESSMENT MODEL

1. Model Purpose and Capability

The Airframe Depot Readiness Assessment Model (ADRAM) represents approximately $738M of the $1.16B fiscal year 2005 Aviation Depot Maintenance Program. Existing in a Microsoft Access database, the model was designed to forecast the dollar cost to achieve a prescribed level of fleet readiness. The output is the total cost to fund the number of airframe inductions required to achieve the targeted readiness level. The ADRAM can also be used to determine the level of airframe readiness achievable, given a certain level of budgetary constraints. These funding drills can be simulated for each type of airframe incorporated in the Aviation Depot Maintenance program. It can estimate the impact of various resource allocation decisions by Type/Model/Series (TMS) and by squadron.

The simulated output is dependent upon the respective assumptions pertaining to the number of airframe inspections, rework, and emergent repair requirements of Navy aircraft that will be scheduled in order to meet the established CNO readiness goals. These scheduled events are primarily based on elapsed time. Each airframe has a certain time interval in which the airframe is required to be scheduled for maintenance or inspection in accordance with the Intermediate Maintenance Concept (IMC). The intention behind the IMC construct is to improve airframe readiness at reduced operating and support costs by planning on an interval basis, more frequent depot maintenance with smaller work packages. According to the fiscal year 2005 budget estimate, the E-2, E-6, F/A-18, H-1, H-53, H-60, P-3, and S-3 aircrafts are incorporated under the IMC concept. Another assumption, however, impacts the output of the ADRAM. There are induction requirements that arise from failures identified during inspections. NAVAIR predicts the number of airframe failures expected within the year of consideration and that number becomes an input into the ADRAM. The primary inspection for the ADRAM is the Aircraft Service Period Adjustment (ASPA) Inspection. Under the ASPA concept, airframes are inspected to determine if their service life can be extended another twelve months. Only airframes whose material conditions prevent the end date adjustment are inducted into the depot for Standard Depot Level Maintenance (SDLM). Based on these underlying assumptions and the key drivers of the airframe model, the ADRAM has the capability of estimating the performance or readiness levels required for each aircraft Type/Model/Series (TMS) in order to achieve pre-established CNO airframe availability goals.

2. Model Drivers and Components

One of the key performance parameters for performance-based pricing models is for the simulated output to be linked to the Chief of Naval Operation's goals. For the aviation programs, the CNO has pre-established readiness goals listed in the Aircraft Program Data File (APDF). From the APDF, the Primary Aircraft Authorization (PAA) can be determined. PAA is the targeted number of operational aircraft the CNO wants available for each squadron. Current aircraft objectives are to achieve 100 percent of PAA for deployed squadrons and 90 percent of PAA for non-deployed squadrons (DON, 2003b). As depicted in the top half of Figure 3.3, the Type Commanders provide the deployment percentages so that the CNO's readiness goals can be determined for program planning. PAA is one of the key cost drivers in the ADRAM. The greater the disparity between required availability and current operational aircraft in inventory, the higher the costs will be for budget considerations.

The other key cost driver for the ADRAM is Depot turnaround time (TAT). TAT is the total amount of time that elapses from aircraft induction to returning an operational aircraft back to the fleet. TAT is highly dependent upon depot capacity constraints, which are affected by multiple variables. As shown in Figure 3.3, capacity is contingent on the amount of airframe backlogs in the system, the amount of work-in-process carry-ins, and the depot core minimums which are constrained by law. United States Code

Title 10, Section 2466 stipulates the at least 50 percent of the funds made available for depot maintenance and repair "shall be used for the performance of depot-level maintenance and repair workload by employees of the Department of Defense" (USC, 2005). All of these data are provided by the depots, owned by NAVAIR, and affects the amount of time it takes to get an airframe repaired through a depot. The slower the TAT, the higher the projected costs will be for Sponsor Program Proposals and POM submissions.

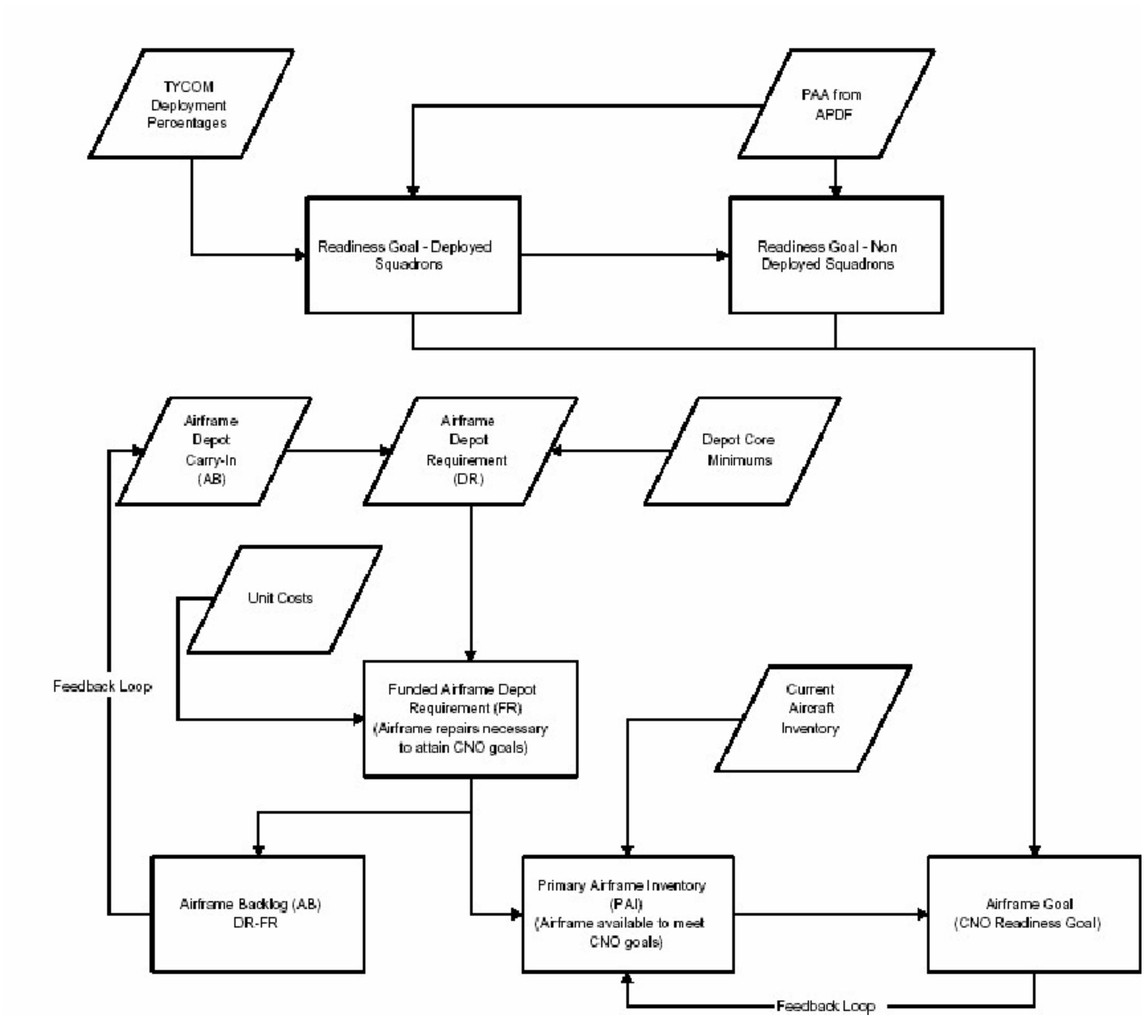

Figure 3.3 ADRAM Flowchart (From: DON, 2003b)

Using all of this information, the backlogs, carry-ins, core minimums, and unit cost by TMS, the airframe funded requirements are determined for each type of airframe. Primary Aircraft Inventory (PAI), which represents airframes available to meet the goal, is calculated using current inventory listed in the Aircraft Inventory AII Exhibit and the backlog and TAT information provided by NAVAIR. PAI is then compared with CNO PAA goals to determine the number of airframes that need to be inducted in order to reach the aviation readiness goal. Current availability is compared with CNO availability goals to determine the performance to goal. This number is multiplied by the repair unit cost per airframe to determine the total level of funding required to accomplish the targeted objective.

3. Costing Methodology

The model generates program cost requirements by inputting standardized cost metrics or depot unit costs for each type of airframe incorporated in the program. These costs per unit (airframe) are provided by NAVAIR. Per the Model Pricing Validation Team (MPVT) documentation, unit costs are determined by using FMB provided inflation rates, revolving fund rates, and material workload standards in organic, inter-service, and commercial depots. The formula for price of each Type/Model/Series is workload standard times the labor rate plus direct material.

D. ENGINE DEPOT READINESS ASSESSMENT MODEL

1. Model Purpose and Capability

The Engine Depot Readiness Assessment Model (EDRAM) defines approximately \$304M of the fiscal year 2005 \$1.16B Navy Aviation Depot Maintenance Program. Created in a Microsoft Access database like the ADRAM, this model is designed to forecast the total cost of meeting an engine readiness goal. The output of the model is the cost associated with a predicted amount of engine inductions necessary to return aircraft engines to Ready-for-Issue (RFI) status to achieve CNO objectives. Similar to the ADRAM, the EDRAM can also reveal program impacts of various resource funding levels. The model can be used to estimate the number of engines that can be inducted based on a certain level of funding and it can drill down through each Type/ Model/ Series.

2. Model Drivers and Components

The current CNO goal for engine requirements is to have 90 percent RFI spares status and zero bare firewalls (DON, 2003b). This means that for the total aircraft inventory contained in the Aircraft Inventory AII Exhibit, each aircraft is required to have the respective number of engines installed that the aircraft is designed to have and the fleet is to be supported with at least a 90 percent spares pool.

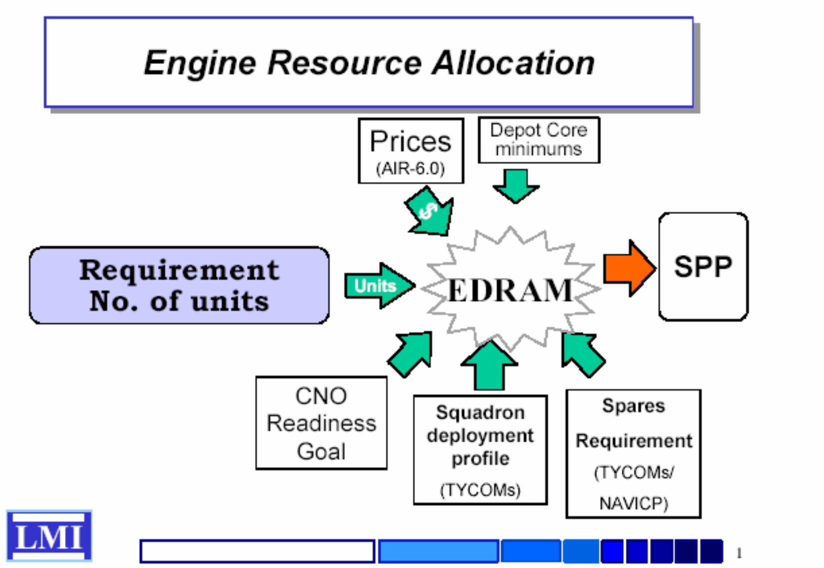

Figure 3.4 EDRAM Components (From: DON 2003b)

As depicted in Figure 3.4, TYCOM provides the deployment data and the Ready-for-Issue requirements that are used to calculate the CNO program readiness goals.

Four key components drive the costs of the EDRAM projections. They are the Intermediate Maintenance Activity (IMA) production, the Flying Hours and Mean Engine Flight Hours Between Removal (MEFHBR), and the Aircraft Inventory Exhibit AII mentioned above. The IMA production is based on assumptions made by N781, as to the level of capacity the Intermediate Maintenance Activities will have to meet engine demand. Only those engines that cannot be repaired at the IMA level are inducted at the Depot Maintenance level. The more engines assumed to need Depot level maintenance, the higher the budget request for the EDRAM. The closer the assumption is that the

IMAs have the appropriate capacity to manage demand, the lower the program costs will be in the SPP.

Engine usage is another key driver in predicting induction requirements. This is generally based on the number of forecasted flying hours listed in the budget OP-20 exhibit and the MEFHBR. The engine removals for each TMS are calculated using the most recent 12 month average of MEFHBR and dividing that into current projected flying hours. The remaining calculations are very similar to the ADRAM forecasting procedures. Engine requirements, based on backlog and depot core minimum data are determined. Comparing primary engine availability with the targeted engine availability generates the number of engine inductions to perform to the CNO's goal. NAVAIR provides the unit costs per engine repair and the EDRAM calculates the total engine program costs to perform at a certain level of engine readiness.

3. Costing Methodology

The costing methodology employed for the Engine Depot Readiness Assessment Model mirrors the process discussed above for the ADRAM. Inflation rates, revolving fund rates, and workload standards are all considered in deriving unit costs. The price for the Navy Aviation Maintenance Program to induct an engine for each TMS is obtained by multiplying the labor rate by the workload standard and adding the direct material.

E. NUCLEAR PROPULSION TECHNICAL LOGISTICS (NPTL) MODEL

1. Model Purpose and Capability

The Nuclear Propulsion Technical Logistics (NPTL) Model represents 19 percent, or \$191M of the \$1.03B appropriated to the 1B5B Ship Depot Operations Support budget category for fiscal year 2005. The model was developed in Microsoft Excel to calculate the costs of achieving various levels of performance within an acceptable, prescribed risk level. Specifically, the NPTL calculates man-years of effort required to perform functions based on depot maintenance availability schedules, reactor plant types, nuclear ship force levels, and average reactor plant age by fiscal year. Calculated man-years are converted into a budgetary figure for POM and budget development efforts. The model enables program managers to perform gap analysis on the required level of man-years as simulated by the NPTL and compare it to the man-years currently available. This drill is executed and costed at various performance levels, with particular attention being focused on the minimal level of performance that can be planned that still fits within the acceptable risk level of workload accomplishment.

2. Model Drivers and Components

There are four key components driving the costs of the Nuclear Propulsion Technical Logistics Model. As pictured in Figure 3.5, the output of the NPTL is dependent on the nuclear fleet age, the time spent in nuclear availabilities, the nuclear force level, and the number of different nuclear core types that need to be maintained and supported.

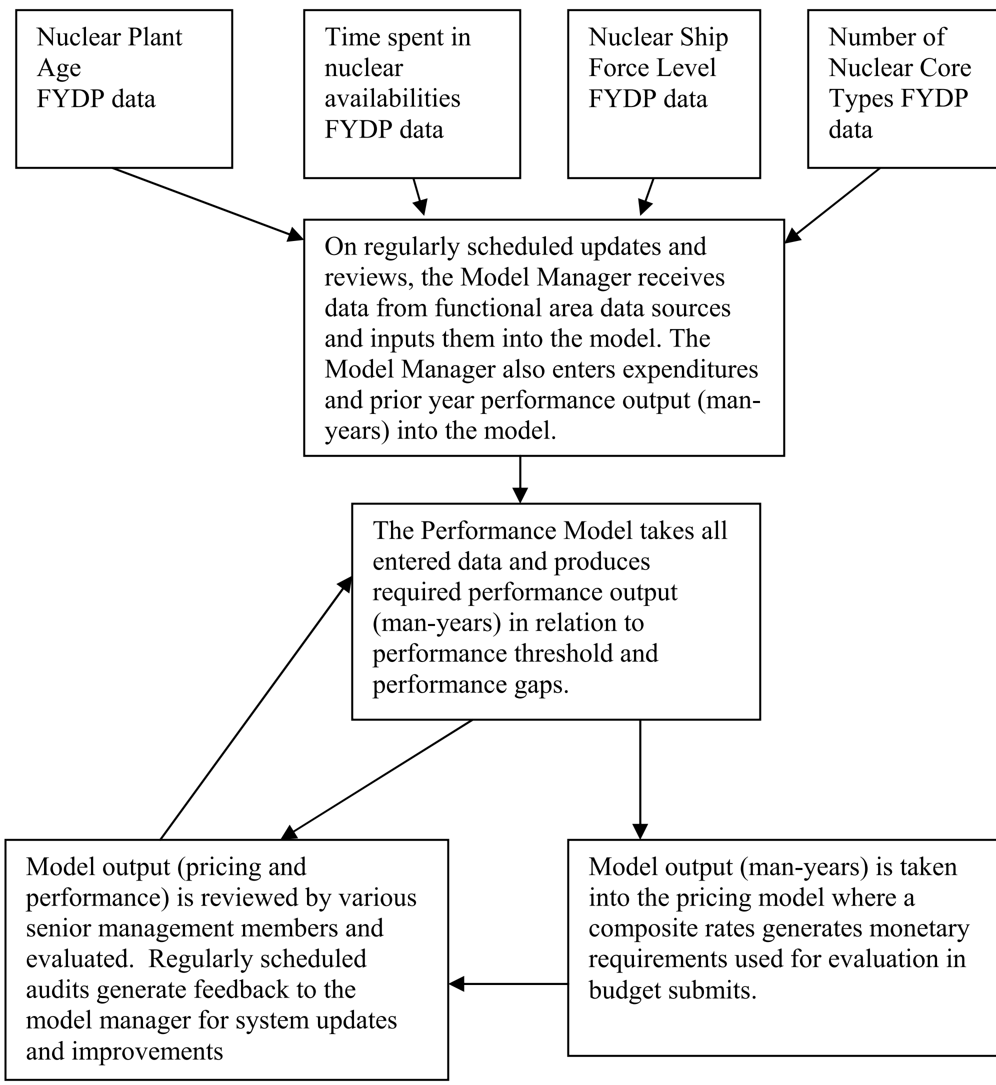

Figure 3.5 NPTL Flowchart (From: DON 2004c)

The nuclear fleet age functional area comprises 38 percent of the total NPTL costs. Data for determining the average ages of the nuclear plants in the fleet comes directly from the Naval Nuclear Propulsion Plant Aging and Ship Life Extension Annual Assessment Memorandum. Increases in the average age of the plants typically result in increases in maintenance requirements and technical support. Decreases in plant age, through ship inactivation decisions, decreases the costs of the NPTL program. Model output is affected by changes in the activation and inactivation plan throughout the Future Years Defense Plan (FYDP).

Comprising 28 percent of the NPTL costs, the nuclear ship force level monitors every activated nuclear powered ship and the number of ships within each class. Force level data are extracted from the Ships and Aircraft Supplemental Data Table (SASDAT) which is maintained in the PBIS database. Increases in force level generate an increase in overall program costs.

An additional 28 percent of the Nuclear Logistics Program costs emanate from the time each plant spends in nuclear availabilities. Engineering and maintenance support costs increase as the time ships spend in nuclear availabilities increase. The fewer the number of availabilities in a given year, the lower the NPTL program costs should be for POM and budget development. The annual nuclear availability schedule is provided by the Fleet Modernization Program Information System (FMPMIS) and the model can be updated as the FMPMIS updates are issued.

The remaining 6 percent of NPTL costs are incurred by the number of different nuclear core types available in the fleet. The model manager updates this functional area manually at each model review. Costs for this area are positively correlated with the number of separate types of nuclear plants. The greater the diversity in nuclear plant types, the greater the requirement for increased technical and engineering expertise. As the scope of technical support increases, the costs of the program increase proportionately.

3. Costing Methodology

The Nuclear Propulsion Technical Logistics Model calculates the required level of man-years necessary to execute tasks concerning the four functional areas discussed above. Once the man-years (performance level) are determined the pricing component in the model converts the man-years to a cost figure. The total cost per man-year is multiplied by the total man-years to develop the funding level to perform within a specified risk level of workload accomplishment. The source of the cost per man-year and how it is calculated has not been included in the NPTL model documentation.

F. ENVIRONMENTAL RESTORATION MODEL

1. Model Purpose

The Environmental Restoration Program performance model comprises an annual budget of approximately \$250 to \$300M (DON, 2004d). Designed in an Oracle database, the model assists the Navy in complying with local, state, and federal laws and regulations pertaining to hazardous waste cleanup and military munitions remediation. Specifically, the Restoration model forecasts the total environmental restoration project requirements and estimates the costs associated with the ensemble of projects. First, the model generates the maximum performance level (PL-MAX) to fully comply with all environmental requirements and provides the estimated cost of achieving this goal. Then the model determines the minimum performance level (PL-MIN) that can be accomplished given certain level of funding that satisfies all legal and regulatory requirements. This allows the Navy to determine the basic level of funding required by the Environmental Restoration program and to understand the legal ramifications of not funding to that level of performance. This model has the capability of predicting the requirements for and prices associated with individual projects as well. High risk sites can be analyzed separately and tracked for completion in a particular year. For a given amount of resourcing, the model can be used to predict the percentage of project completion in high risk projects. Medium and low risk projects can be evaluated independently and tracked for costs and schedule completion during the life of the project. Performance is measured by the percentage of high, medium, and low risk project sites that were completed in the scheduled year as specified in the Defense Management Goals.

2. Model Drivers and Components

Three main drivers control the performance output of the Restoration model. Legal and regulatory drivers, DOD and DON program goals and funding guidance, and the condition of the contaminated sites being addressed impact performance requirements simulated by the model. As depicted in Figure 3.6, the Environmental Restoration Model is a conglomeration of interrelated business processes and systems, each being influenced by the key model drivers to determine the appropriate level of funding to meet targeted performance objectives.

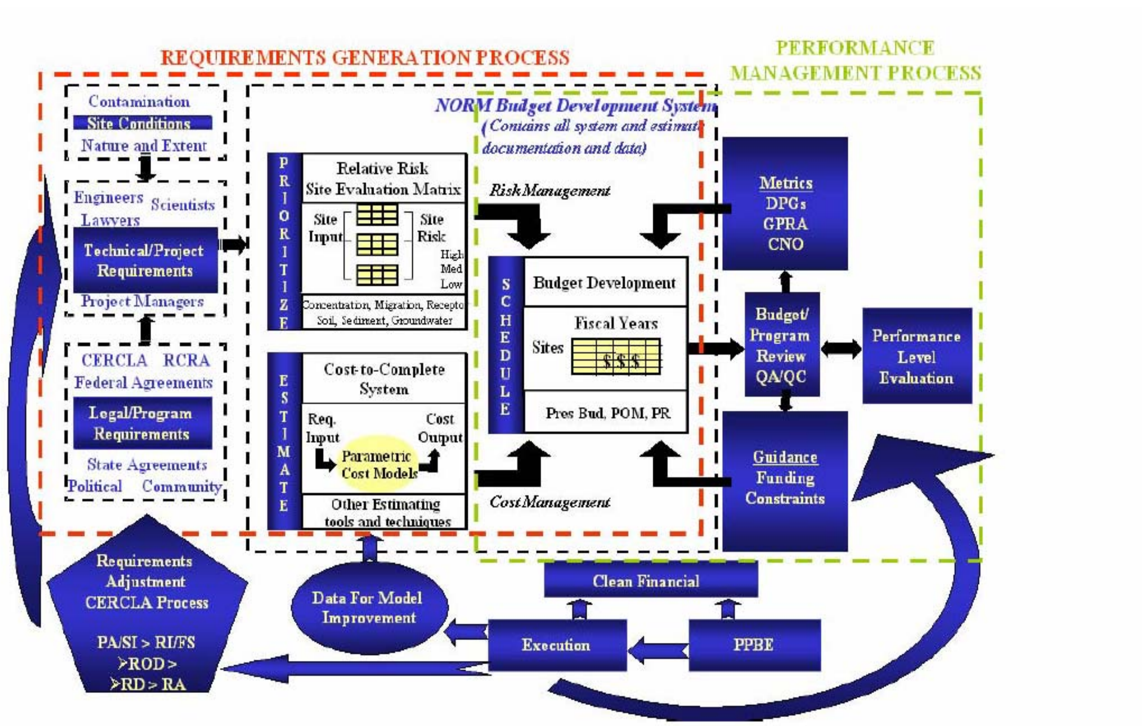

Figure 3.6 Environmental Restoration Model Components (From: DON 2004d)

The Requirements Generation process is composed of the legal and site drivers, a risk evaluation primer, and a cost-to-complete parametric estimating system that feeds into the Normalization of Data (NORM) Budget Development and Management System.

The legal drivers derive from the hazardous waste cleanup and control requirements of the Comprehensive Environmental Response, Compensation and Liability Act (CERCLA) and the Resource Conservatory and Recovery Act (RCRA). Based on the these laws, as well as DOD and local guidance, subject matter experts on integrated project teams conduct site assessments to determine the hazardous condition of areas of interest. These groups of engineers, scientists, and lawyers estimate a particular project's requirements and provide an educated estimate of the applicable costs associated with a project. Next, complying with DOD risk management guidance to evaluate environmental projects by measuring the relative risk of threats to human health and the environment, the integrated team categorizes the projects into high risk, medium risk, and low risk segments. The various performance levels are calculated using the parametric cost estimating system in the cost-to-complete component. The performance level to implement and/or complete all projects known at that time and to comply with all legal restraints are priced without any fiscal constraints and fed into the NORM budget development system.

The performance management process entails the incorporation of the primary model cost driver, the established Defense Environmental Restoration Program (DERP) performance metrics and goals. The Office of the Secretary of Defense (OSD) established environmental program goals and objectives in the Defense Management Goals. They are to complete (DON, 2004d, p. 11):

- 50% of the identified high relative-risk sites by the end of FY 2002
- 100% of the identified high relative-risk sites by the end of FY 2007
- 100% of the identified medium relative-risk by the end of FY 2011
- 100% of the identified low relative-risk sites by the end of FY 2014

This guidance, along with the OPNAV instruction 5090.1B detailing Navy compliance requirements with environmental laws and regulations are used for the development of program performance levels within fiscal constraints. PL-MAX and PL-MIN levels are calculated by the Environmental Model to allow senior leadership to review the critical performance requirements that need to be funded and to understand the legal ramifications and environmental impact of resource allocation decisions.

3. Costing Methodology

As shown in Figure 3.6, the Environmental Restoration Model contains a pricing component, the cost-to-complete (CTC) system, based on parametric cost estimating relationships. The CTC identifies component parameters and determines the costs driving those parameters to establish cost estimating relationships. As shown in figure 3.7, this costing tool has the capability of estimating direct costs, integrated cost estimating relationships, and also of incorporating expert-defined Government cost estimates (PwC, 2001). Using multiple regression techniques and multivariate equations and activity-based interactions, the CTC provides costing data for specific site work, professional labor activities, sampling tasks, and allows for the costing of additional user requirements on selected projects.

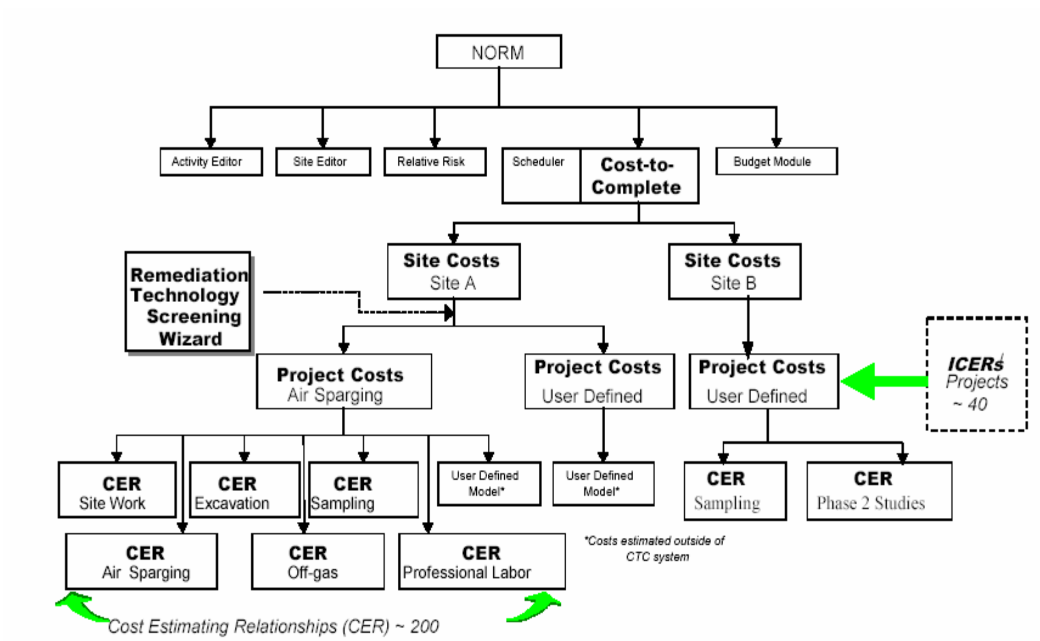

Figure 3.7 NORM CERs (From: PwC, 2001)

The model's cost estimates are only as accurate as the source of the information used to populate the various cost elements. The cost data compiled from various sources were validated and accredited by PricewaterhouseCoopers in 2001. Each cost element used in the model is "codified in the Environmental Cost Element Structure…the national standard classification sponsored by the Building Economics Subcommittee of the American Society of Testing and Measurement (ASTM) (PwC, 2001, p. 12).

G. SUMMARY

Performance-based pricing models currently define the requirements for over half of the Total Obligation Authority (TOA) in the Department of the Navy. Through the Sea Enterprise BOD program and budget development process, the Navy is making progress towards integrating performance models throughout each phase of the planning, programming, budgeting, and execution (PPBE) process. Based on interviews with the program managers of the models, the Aviation Depot Maintenance Model, the Nuclear Propulsion Technical Logistics Model, and the Environmental Restoration Model were incorporated in the six phases of POM development outlined in section B of this chapter. In fact, for POM-06, each of these performance models received the exact amount of money estimated by the model to execute to a certain capability level. Based on the fundamental assumptions and parameters constructing the model, program requirements were simulated and priced for Sponsor Capability and Program Proposals. The three fully accredited models were presumed to be credible and were rewarded with a fully funded program. The Environmental Restoration Model received a budget increment for its program to support additional cleanup efforts on the island of Vieques, PR. This supports an underlying vision in the President's Management Agenda, to fiscally reward those high performing programs that have successfully integrated performance information with budgetary resources.

Performance models played a significant role in the planning and programming phase. Yet, in the budget development phase, these models were not used by FMB directly. In fact, none of the FMB points of contact had or has access to the models reviewed for this research. Often, when making resource allocation decisions, FMB will have discussions with program managers to determine the impact of proposed budget cuts and to make tradeoff decisions. However, the integration of performance-based pricing models in the budget development phase can be improved. In the quest to achieve the vision of the President's Management Agenda and the Budget and Performance Integration Initiatives, the process appears more like performance-based programming than performance-based budgeting.

In the execution phase, based on interviews for PR 2005, there appears to be less emphasis on tracking program performance and feeding information back into the models than on planning and programming for the next Program Review. Program managers are the primary users of the models and they have the propensity to be more forward-focused in their performance efforts.

The chapter has discussed the role of performance-based pricing models in the Department of the Navy's POM and budget formulation process. It has also addressed the key model drivers and components of the Navy's three fully accredited performance models and how they generate program requirements and capability levels. Each model's costing methodology was evaluated to determine how the models develop budget estimates for simulated performance levels. The next chapter will analyze the verification, validation, and accreditation process employed on these performance models to ascertain the level of accuracy and confidence these models were built to achieve.

THIS PAGE INTENTIONALLY LEFT BLANK

IV. CURRENT DON VERIFICATION VALIDATION & ACCREDITATION (VV&A) TECHNIQUES

This chapter examines the verification, validation, and accreditation techniques and processes employed on the Navy's three performance-based pricing models accredited with the highest user confidence. The only performance models to reach that distinction to date are the Aviation Depot Maintenance Model, the Nuclear Propulsion Technical Logistics Model (NPTL), and the Environmental Restoration Model. The goal of this chapter is to compare the VV&A tasks performed on these models and pattern-match them with the activities commonly practiced in industry. The chapter concludes with a summary of the analysis of the VV&A process practiced by the DON to certify the use of these performance-based pricing models for programmatic and budget development decisions.

A. INTRODUCTION

"All models are wrong. Some are useful" (Carson, 2002, p.52). This contention of the world-renown statistician, George Box, underlies a fundamental truism when attempting to develop models to simulate reality. In the execution of any project, like building a performance pricing model to predict resource requirements, project and program managers have to negotiate and make tradeoff decisions within a triple constraint of affordable costs, schedule, and performance parameters (DOD, 2003c). The higher the level of performance required, the higher the project's cost will be. Working under a constrained or compressed schedule significantly impacts the total costs and performance attributes attainable as well. Yet, even if expense was no matter, when developing reality-based models for complex systems or processes, the model's performance accuracy can never be 100 percent. Law and Kelton maintain:

> A simulation model of a complex system can only be an approximation to the actual system, regardless of how much effort is put into developing the model. There is no such thing as an absolutely valid [simulation] model. The more time (and hence money) is spent on model development, the more valid the model should be in general. However, the most valid model is not necessarily the most cost-effective one. For example, increasing the validity of a model beyond a certain level may be quite

> expensive, since extensive data collection may be required… Furthermore, we question whether hypothesis tests, as compared with constructing confidence intervals for differences, are even the appropriate statistical approach. Since the model is only an approximation to the actual system, a null hypothesis that the system and the model are the "same" is clearly false. We believe that it is more useful to ask whether or not the differences between the system and the model are significant enough to affect any conclusions derived from the model (Page et al, 1997, p. 396).

The verification, validation, and accreditation efforts are designed to build a feeling of confidence for model users so that they may make informed decisions based on reliable data. In performance-based pricing models, the users are making resource allocation decisions and justifying budget submission requests based on the output of the models, often for large dollar value or safety-related programs. Therefore it is essential that these performance models generate accurate information. "Confidence in a particular model or simulation must be justified before its results are used to make decisions involving large sums of money or risk to human life" (DON, 1999). Modeling and simulation (M&S) involves the development and implementation of a synthetic and logical representation of a system or process to understand and observe predictable behaviors based on real world phenomena. Verification agents are tasked to ensure the M&S is built correctly based on the articulated concept and design. Validation agents test whether the right model was built. They check to see if the model demonstrates a sufficient representation of reality. Finally, when the V&V effort is complete, the accreditation process certifies the modeling and simulation (M&S) tool for its intended use. The entire process is designed to mitigate the risks associated with using performance and requirements models for critical decision-making.

The Department of the Navy VV&A process was established in 1999 by the Secretary of the Navy Instruction 5200.40. Applicable to all models in the DON, the various steps involved in V&V were decomposed into four primary events collimated with the model development phases throughout the M&S lifecycle. The process begins with the validation of the developer's conceptualization of the intended application and ends with the certification that the output of the M&S accurately and reliably represents the known behaviors of the modeled program or system. Model managers were directed to use the V&V templates established by the Office of the Chief of Naval Operations to adequately provide the necessary information for the Director, Assessment Division (N81) to ascertain the level of accreditation to ascribe to a performance model (DON, 2003a). The acceptability criteria and accreditation decision options were also established in the OPNAV Memorandum for "Accreditation and Use of Performance/Pricing Models in POM-06" issued in September 2003.

In 2002, the Navy Modeling and Simulation Management Office (NAVMSMO) was established in support of the Navy's modeling and simulation vision, to use "world-class models, simulations and simulators as tools to meet the future challenges of supporting force readiness, structure, and transformations" (DON, 2002, p. 2). The role of NAVMSMO is to provide procedural guidance and serve as a central repository for all matters involving modeling and simulation applications in the Navy. In 2004, NAVMSMO issued more detailed guidance to better align VV&A practices with those typically employed in the business modeling community (see Figures 2.3 and 2.4 in Chapter 2) and DOD. The primary difference is the NAVMSMO version condenses the implementation and integration phases and performs results validation separately on the finished product. The general process, however, is very similar. Both processes emphasize the importance of beginning the VV&A process at the very beginning of the model development project and continually conducting V&V "throughout the M&S process whether it is a new development or a legacy M&S undergoing modification" (DON, 2004b). VV&A activities should be paralleled to each phase of model development.

The following section analyzes the DON procedures used on the fully accredited performance models and provides recommendations based on the VV&A techniques and activities typified in the modeling and simulation community and the Defense Department. First, a description of what the modeling and simulation phase and V&V process entails is discussed. Then the V&V procedures used on the three fully accredited performance models are analyzed based on current OPNAV techniques, as delineated in Appendix A. Last, recommendations are provided upon pattern-matching the OPNAV techniques with 2004 NAVMSMO guidance and the critical tasks practiced in DOD and the business modeling community (Figure 2.4).

1. Requirements Phase

a. Description

The requirements phase of model development entails the clear and unequivocal definition of the need for the model through requirements elicitation and analysis. The "most critical piece of the M&S development and V&V activities falls in the very beginning of the life cycle" (Chew and Sullivan, 2000, p. 814). Complete information pertaining to the specific performance and quality expectations of the model user are thoroughly described in written form and provided to the model developer. The documentation details the key performance parameters, the level of fidelity, and capability requirements critical for acceptability and delimits those that are lower priority desirables. Quality requirements, traceability requirements, configuration management requirements, and user interface requirements are all included in the documentation to provide the foundation for the M&S development effort and V&V functional events.

Ensuring the full understanding and analysis of the user needs and requirements before any work is performed for model creation, maximizes the opportunity for model developers to deliver the right model suitable for the right application within the right resource constraints. As the model engineers assess and analyze the stated requirements, they document the potential inconsistencies and risk areas and resolve the discrepancies with the model user before time and resources are spent building the model. Identifying problem areas and misunderstandings early in the process enhances the possibility of achieving significant cost savings over the life cycle of the project. Finding a latent mistake after system completion may force the model development team to reconstruct the model and begin the entire evolution all over again. To mitigate the risks of this happenstance and ensure the model development project team has the proper information to successfully meet user needs, the verification and validation agents evaluate the requirements documentation for consistency, clarity, testability, and suitability to model purpose.

b. Analysis of Practice

While the DON guidance discusses the need to "specify and analyze requirements" (DON, 1999, p.3), current practices do not incorporate requirements V&V into the VV&A standard processes. The models being developed for informing resource allocation decisions and budget formulation and implementation are verified and validated using the performance pricing model V&V template (or some version of it) depicted in Appendix A. As evidenced by the template, the requirements analysis is not the beginning of the credibility building methodology of Navy VV&A.

As mentioned in Chapter III, the Aviation Depot Maintenance Model actually consists of two separate models. There is the Airframe Depot Readiness Assessment Model (ADRAM) and the Engine Depot Readiness Assessment Model (EDRAM). The ADRAM and EDRAM verification and validation process commenced years after the models were fully developed using an earlier version of the standard V&V template depicted in Appendix A. The requirements model was developed in 1999 by the Logistics Management Institute (LMI) and has been used in every Program Objective Memorandum (POM) cycle since Program Review for fiscal year 2001. Based on the V&V documentation, the ADRAM/EDRAM is a Microsoft access based, legacy model that had not gone through a rigorous VV&A process. Instead they were "vetted through Navy Office of Budget (FMB) and the Office of Secretary of Defense, Program Analysis and Evaluation (OSD, PA&E)" (DON, 2003b, p. 6). A few user friendly upgrades were made to the model in 2003 along with the model's initiation into the verification and validation process. The existing modeling and simulation requirements and program performance measures for the current application of the Aviation model were not reviewed and documented for accreditation purposes.

The second performance model under review, the Nuclear Propulsion Technical Logistics Model (NPTL) also followed the V&V template. The performance model was verified and validated without documenting the requirements determination and evaluation process. The model was developed in 2003 and verified and validated in 2004 without the explicit delineation of model requirements and user needs.

Ascribing to the same V&V methodology as the other models, the Environmental Restoration Model team did not include requirements V&V into their technical analysis processes either. The Restoration Model is a legacy model that has been extensively used for approximately ten years prior to its implementation into the VV&A process. Based on the satisfactory and proven results of the modeled program in budget development over those years, the V&V team did not focus their efforts on the existing requirements of the model.

c. Recommendations

The Aviation Depot Maintenance Model, the Nuclear Propulsion Technical Logistics Model, and the Environmental Restoration Model all executed the verification and validation processes of their program models without performing the essential V&V activities for the requirements phase. The reason for this omission was the fact that the internal guidance distributed via the verification and validation templates pretermits the requirement to do so.

Recommendation 1. Based on updated DON guidance in the Navy Modeling and Simulation Verification, Validation, and Accreditation Handbook and the critical V&V techniques practiced by the modeling and simulation community (Figure 2.4), it is recommended that the V&V template be modified to incorporate the tasks associated with requirements V&V. This involves the derivation of VV&A plans to drive the modeling and simulation validation process. The accreditation plan identifies the acceptability criteria by which the M&S will be evaluated. These criteria, sometimes referred to as exit criteria, are quantitative and qualitative measures of effectiveness, performance, and suitability by which each M&S requirement must be addressed. Providing the standards by which the performance model will be assessed enhances the quality of the verification and validation planning effort. "V&V planning identifies tasks that address the acceptability criteria, M&S requirements, resources, and timelines" (DON, 2004). It provides the V&V tasks and methodology to be exercised, specifically tailored for the needs and use of the performance model. The plan implements a schedule of activities, determines the appropriate level of funding needed, and identifies potential risks of the modeling project. V&V plans for legacy models, like the Aviation Depot

Maintenance Model and the Environmental Restoration Model will include documentation on the development history, previous V&V information, and information on intended modifications and enhancements to the existing performance model (DON, 2004).

Requirements V&V basically entails reviewing all documentation on user needs, the intended use of the model, the key performance parameters, configuration management, and expected quality of the model. This phase ensures the requirements are clearly written, "consistent, testable, and complete" (Chew and Sullivan, 2000, 815). Future performance modeling and simulation efforts should include requirements verification and validation in their VV&A methodologies to enhance model credibility.

2. Conceptual Model Phase

a. Description

Using the information generated from the requirements V&V, the model developer's conceptualization of the need is thoroughly defined. "A conceptual model typically consists of a description of how the M&S requirements are broken down into component pieces, how those pieces fit together and interact, and how they work together to meet the requirements specified" (Chew and Sullivan, 2000, p. 815). The goal is to present the model developer's intended approach or framework to designing the model considering user requirements, intended use, acceptability criteria, and data availability. All factors from model assumptions, algorithms, strengths, and weaknesses to cost constraints, risks, and key performance parameters are examined and addressed. Discrepancies or incongruities identified during this phase of V&V result in a modification to the conceptual model, stated requirements, or acceptability criteria. Concept refinement serves as the vital link between the requirements definition and the model design phase.

Validating the fidelity of the conceptual model and verifying its alignment to the user requirements and the acceptability criteria is one the most critical phases of the VV&A effort. Identifying areas of impeded interoperability and the availability of critical data elements early in the modeling and simulation process allows for less expensive architectural corrections and the necessary reengineering of system requirements. Also, providing V&V documentation of the conceptualized model, clearly defining how the model works and interacts to generate the requirements builds user confidence in the model.

b. Analysis of Practice

One can see in the Conceptual Validation section in the Navy template in Appendix A, the Navy requires the accomplishment of most of the critical tasks associated with the V&V of the conceptual model. The template includes the requirement for a graphical representation and written description of the conceptual framework as well as key drivers and their importance to the functionality of the model. Key drivers detail the various assumptions, policy guidance, and data sources applicable to the desired performance of the model. Attached to the V&V template is a section for the key drivers, or data sources to be verified and validated. This portion covers the analysis of the availability of the data, the risks, weaknesses, and limitations of the data, and the methods of collection.

Conceptual validation also requires the decomposition of the model components and the costing methodology used for price estimating. Next, the description of the output with an explanation as to how the output relates to the model's primary objective is addressed. The metrics devised to evaluate the output effectiveness are then expected to be benchmarked against "industry standards or other accepted standard" (Appendix A). The template covers all of the activities necessary for the conceptual phase except the documented assessment of the model's suitability to meet the acceptability criteria and its linkage to user requirements.

The Aviation Depot Maintenance Model (ADRAM and EDRAM) concisely answered some of the requirements necessary for the conceptual validation process. A graphical framework of the conceptualized model was provided with an explanation of the process. The majority of the data required to calculate the aviation requirements were owned and maintained by sources external to the program. These sources were identified and briefly explained on the Data V&V attachments. The availability and V&V of the data was not part of the documentation, however. Also, the ADRAM and EDRAM models are significantly influenced by the output of the Flying

Hours Program (FHP) Model. VV&A documentation from the FHP was not provided to validate the use of this input to the model. And lastly, the Aviation Depot Maintenance Model V&V team stated that there were no industry standards to benchmark program metrics against. There were no similar civilian-equivalent programs to compare performance measures.

The conceptual validation effort for the Nuclear Propulsion Technical Logistics (NPTL) model focused mainly on the key model drivers and components. Without presenting a graphical representation of the developer's interpretation of the performance model and its underlying functionality, the V&V team provided a written description of the components and how they relate to model output. Next, using the data V&V forms, they explained the availability and limitations of the data. The sources were identified and evaluated by the V&V team as being highly credible with low risks. However, the V&V team did not completely address many of the tasks expected during conceptual validation. The documentation did not reveal the model assumptions, equations, or algorithms associated with generating the output. A detailed description of the component and subcomponent interactions and the costing methodology used to determine the funding requirements were not exemplified. As well, the performance metrics to benchmark model performance against industry or other acceptable standards were not mentioned in the report.

The most thorough V&V analysis for the conceptual modeling phase was conducted by the Environmental Restoration Model Team. Most of the requirements listed in the V&V template were responded to in detail. The conceptual model showed the component and subcomponent interrelationships and how data input gets transformed into desired outputs. A clear definition of what the output of the model is and the desired program goals and their linkage to the user requirements was described. Then the key model drivers and data sources were discussed, verified and validated. The V&V team clearly stated the assumptions, policies, and constraints applicable to the performance of the model, as well as legal restrictions that influence funding decisions. Attached to the report was the Accreditation report and recommendation for the Cost-To-Complete model, a major component of the Restoration program that feeds into the model. This documentation included the algorithms, equations, and cost estimating relationships embedded in the performance model. The V&V team was unable to find a similar program in industry or the Department to benchmark metrics against but does provide information on how the Environmental team works with industry and federal agencies in data sharing, experiences, and technologies. The only area not addressed, and not currently requested in the V&V template, is the analysis of how the model meets the accreditation criteria. Making this assessment early prevents the expensive adjustments that may ensue later in the modeling and simulation process. For legacy M&S, the assessments of the new or modified requirements and how they relate to the acceptability criteria is equally essential.

c. Recommendations

There are a couple recommendations for this V&V phase.

Recommendation 2. First, it is recommended that the acceptability criteria be assessed for model fidelity and that the V&V team ensures that the conceptual model framework traces to these criteria (Chew and Sullivan, 2000). Evaluating this linkage at this stage significantly increases the likelihood that the performance model will meet the acceptability criteria after the development is complete.

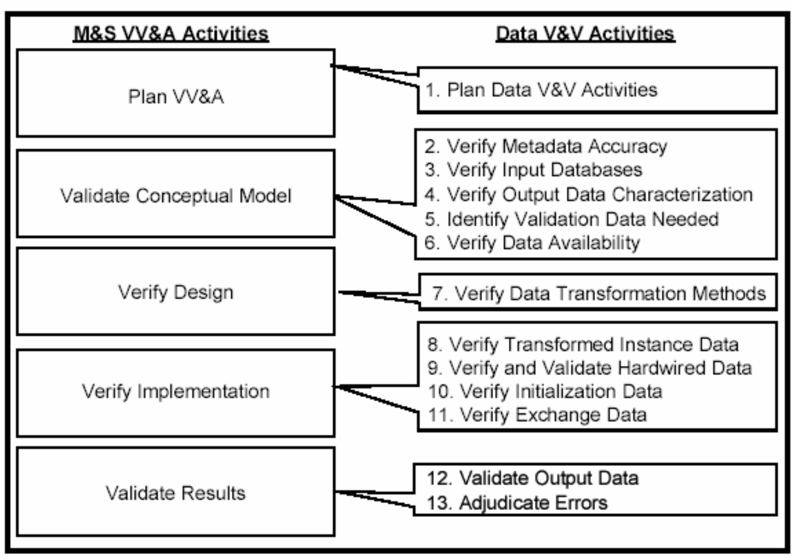

Figure 4.1 Data V&V Activities (From: DON, 2004b)

Recommendation 3. Based on 2004 Navy guidance from the DON M&S VV&A Implementation Handbook, it is recommended that Data V&V be spread along five categories as depicted in the Figure 4.1 above. Instead of performing the data V&V all in the conceptual model validation phase, there should be data V&V activities involved in each stage of the model development process. "By stepping through the data V&V activities shown…in parallel with V&V implementation activities, the V&V agent can ensure that the proper data was identified, obtained, and used properly in the M&S" (DON, 2004b, p. 15). Regardless of how well built and architecturally sound the model may be, the M&S product is only as good as the quality of the data and information obtained by using it.

The Navy requisites currently in place, based on the template, are sufficiently thorough in scope for the verification and validation phase of conceptual model development. Ensuring the requisites are met is as essential to the credibility building process as is establishing the exit criteria for each phase. Whether it is a new development or a legacy model with enhancements or modifications to the existing program requirements, sufficient V&V along with thorough documentation builds user confidence in the model, assures the model's ability to meet current objectives, and improves the learning curve for follow on performance modeling efforts.

3. Design Phase

a. Description

"After the conceptual model is verified and validated, the developer produces a detailed design that describes exactly how the conceptual model will be coded or fabricated" (Chew and Sullivan, 2000, p. 816). All of the elements and components outlined in the conceptual model are incorporated into a design that shows how the various factors will be developed. Software coding is identified and analyzed. The structural, system design is evaluated to determine how the hardware will be developed and how the hardware and software pieces interact and interface with user requirements. The "design features emphasize the functionality, information flow, ordering of processes, and data accessibility" (DON, 1999, p. 4). It is the critical stage that links the developer's conceptualization and user requirements to the system design.

Verification and validation of the design entails assessing whether all of the elements identified in the conceptual model have been consistently and completely translated in the proposed design. It verifies how the features of the hardware and software support the user requirements and how the design framework meets the acceptability criteria.

b. Analysis of Practice

The Navy 2003 template and the 2004 template in Appendix A, reveal the verification requirement of the design phase. The templates show the requisite that a design framework be developed with an accurate description of the design features. Then it adduces the requirement for the V&V team to show how the design links back to the conceptual phase and the user requirements. It does not address the accreditation acceptability criteria or the need to identify the planned software applications and perform hardware emulation. It basically serves a tautological representation of the conceptual model.

The Aviation Depot Maintenance Model team addressed the design verification requirements in two sentences. The documentation stated that the design verification was completed during the model development process. Yet, no graphical depiction of the design was provided. Nor was any written explanation as to how the model was designed and what software, hardware issues addressed. The V&V team provided little useable information for determining how the model was designed, how it linked to the conceptual model, and how it supported the accreditation criteria.

The NPTL model V&V team bypassed the design verification phase altogether. The responses in the template went from discussing the conceptual framework straight to system verification issues (DON, 2004c).

The Environmental Restoration Model, using the 2004 template, documented the V&V activities for the model's design. First, the team provided the same graphical representation that was provided for in the conceptual validation phase. The design model and the conceptual model should not be the same exact representation. The design should define the "components, elements, functions, and specifications that will be used to produce the simulation based on the conceptual model" (Chew and Sullivan, 2000, p. 816). It focuses on the hardware and software that will be used to support the conceptual model. It is a preliminary look at the system before any piece of software code is written or hardware fabricated, if applicable. The V&V team did, however, provide a detailed analysis of how the components and functional processes interact to develop program levels and goals. The report also delineates what those goals were.

c. Recommendations

Recommendation 4. The primary recommendation for the design V&V phase, is for the verification and validation team to require a detailed design with a documented explanation as to how the design links to the conceptual model and user requirements and verify the design linkage to the acceptability criteria prior to entering the next modeling phase. This should be required for both new development and legacy models. For example, the ADRAM and EDRAM models are legacy models. Spending an extensive amount of time and resources on the design phase may be imprudent. However, providing adequate documentation for the design of the models is a credibility issue. Documentation showing the various assumptions and algorithms associated with the design of the model allows the user and the accreditation agent to gain confidence that the model was built soundly. The accreditation agent has to understand how the model was designed to make a determination that the model is suitable for its intended use. Due to the fact that there are only eight acceptability measures critical to the accreditation decision and one of them focuses on the design of the performance model, there should be much more time and effort dedicated to verifying the design and providing essential documentation. If there is no design documentation, they need be to reengineered and recreated to support V&V efforts, reusability, and the accreditation decision (DON, 2004b).

Recommendation 5. Also, it is recommended that the requirement to verify any data transformation methods applicable to the particular model be added to this design verification section of the template (Figure 4.1).

4. System Phase

a. Description

This generally is the stage where the technical solution or model is developed. First, if applicable, the hardware is fabricated and the software code is written. Before these are integrated they are tested and verified. The second stage focuses on system integration, where all of the components and elements of the designed system are desegregated and tested for operational verification.

"System verification is the formal (i.e., documented) test/review process by the M&S proponent responsible for determining that the M&S accurately represents the functional design and has traceability to the conceptual model and the system requirements" (DON, 1999, p. 5). The code and the hardware are verified first to ensure that the accepted design is implemented properly and to identify and resolve discrepancies prior to system integration. Once the system is fully integrated, verification tests whether the user requirements, conceptual model, and detailed design are translated as documented (Chew and Sullivan, 2000).

b. Analysis of Practice

The DON template does a thorough job of capturing essential information for the systems implementation and integration phase. The requirements for this phase include developing a high-level diagram of the performance model and how all of the components and elements interact. The Navy also requests a clear depiction of the cost and performance feedback loops. Responding to the various tasks outlined in the template provides the traceability to the user requirements, to the conceptualized framework, the detailed design, and to the acceptability criteria. The V&V team is required to demonstrate how the model is linked to performance metrics and readiness goals. They must describe how the performance requirements and pricing elements interact for programmatic and budgeting decisions. In addition, the specific test procedures employed to verify the requirements are expected to be included in the V&V report.

The ADRAM and EDRAM V&V teams, guided by an earlier template, provided a cursory explanation of the verification process. A framework of the system was provided for the ADRAM but the framework for the EDRAM was unattainable for this research study. Tests procedures and their corresponding results were not expressed in the V&V documentation. OPNAV also requires the V&V team to state the level of susceptibility of the model's output to fluctuations in the output of other performance models. If the performance model inputs data from the output of another model, this must be clearly stated in the V&V report along with the potential impact on the performance model's outcome by using this input. The Aviation Depot Maintenance Model team identified Primary Aircraft Authorization as the input that would have the greatest impact on model behavior. The V&V team did not expound on the level of susceptibility nor did they include the effect the output from the Flying Hours Program Model would have on the model. The EDRAM listed flying hours and the Intermediate Maintenance Availability as key drivers of susceptibility. That was the end of the system verification and validation effort.

The Nuclear Propulsion Technical Logistics (NPTL) Model Team provided a thorough description of the system verification process. First, the team provided an illustration of the system and an explanation of how each component works and interact to achieve program goals. The system is based on Excel software. The V&V team describes each functional tab applicable to the program and details how the data are manipulated to generate requirements. Constraints to the model are identified as well. The V&V team concludes the phase with a discussion on how the performance model fits with the conceptual model and the design hardware and software components.

The Environmental Restoration Model addresses each task in the Navy template in substantial detail. All of the major components of the Restoration model were individually described and graphically depicted as part of the system. The architectural structure and framework for the Normalization of Data (NORM) component was thoroughly explained and portrayed in a separate illustration. Next, the team detailed the pricing and performance interplay and linkages of the output to the performance goals of the program. The report stated how the output is integrated with the Planning, Programming, Budgeting, and Execution (PPBE) Process. The V&V team demonstrated how the model represented the detailed design and the user requirements for the program.

The test procedure used to verify the system was identified but the results of the test were left out of the report. The results of the model were stated to be susceptible only to the high level of uncertainty inherent in the environmental program.

c. Recommendations

As mentioned above, the DON template covers the vast majority of the required tasks for the system V&V phase. The following recommendations apply to this phase in the V&V process.

Recommendation 6. It is recommended that the tasks required in the template are responded to with thorough explanations and documentation on the particular modeling factor. There is a significant difference in identifying a critical factor and explaining its impact on the planned output. Providing detailed information elicits greater understanding and credibility in the use and accreditation of the model.

Recommendation 7. It is recommended that any tests conducted for system verification be thoroughly described and documented, along with the clear presentation of the test results. Stating that a certain element was tested without explicitly describing the test and documenting the results does not add to the credibility of the model. All acceptance and compliance test results should be clearly stated in the V&V report. "Acceptance testing determines whether all requirements are satisfied… and compliance testing determines if the simulation meets required performance standards" (Chew and Sullivan, 2000, p. 816).

Recommendation 8. Also, per the information in Figure 4.1, the four data V&V activities applicable to this phase are recommended to be included in this section of the VV&A process. They consist of verifying transformed instance data, V&V hardwired data, verifying initialization data and exchange data.

5. Results Phase

a. Description

Once the model has been developed and verified, the validation of the system output is the next critical phase. Results validation is the "formal (i.e., documented) test/ review process that compares the responses of the M&S with known or expected behavior from the subject it represents, in order to ascertain that the M&S

responses are sufficiently accurate for intended uses" (DON, 1999, p. 5). It is the phase where the output of the performance-based pricing model is compared with real world behavior to validate the model's fidelity. There are a wide variety of validation techniques that can be used. There is face validation, benchmarking, stochastic and statistical analysis, historical methods, and subject matter experts (SME) to name a few. As shown in Figure 2.4, there is also depth of testing and breadth of testing metric considerations. Regardless of the approach, performing results validation using "real world (prototype) data increases the confidence in the M&S resources and also provides higher fidelity answers to specific questions" (Caughlin, 2000, p. 876).

b. Analysis of Practice

The DON has stipulated that the results validation process can "only be completed if real world data is available" (Appendix A,). For example, if Program Review 2005 is the first time that a performance model was used, then real world data will not be available until the end of fiscal year 2005. The Navy template also states that full accreditation, which is the highest credibility rating a model can achieve, requires that results validation phase be complete. The verification and validation team is expected to provide documentation showing the results of an actual test using real world data. Any errors and corrections occurring during this phase also need to be fully documented. As part of the required Navy tasks, the V&V team is asked to evaluate, document, and demonstrate how the Navy's accounting system helps or hinders the ability to track the amount of dollars programmed against amounts budgeted and also executed for the particular model. The V&V team then shares any recommended changes to the accounting system. And finally, if subject matter experts were substantially involved in the validation process, their names and credentials should be listed in the documentation.

The Aviation Depot Maintenance Model results validation team tested the model by generating a spreadsheet to calculate the performance requirements and comparing them to the output of the model. This was done for each individual airframe Type Model Series by squadron. The V&V response stated that the M&S results are not and have not been compared to real world data and that "there was no attempt made to reconcile programmatic projections with execution data" (DON, 2003d, p. 6). There are a few issues that arise from this testing procedure. The first is that the test was only focused on the requirement. It did not including the costs associated with the program requirement. For resource allocation decisions, the models need to be able to predict fully costed requirements. The second issue was that the formulas used in forecasting the requirements in the spreadsheet should have been explicitly addressed and related to the formulas used in the Aviation Depot Maintenance Models. If different formulas were used in the spreadsheet as were used in the model, and they predicted different requirements, there needs to be evidentiary documentation to validate that the output of the spreadsheet accurately represents reality. If the same formulas were used in the Aviation Depot Maintenance and the spreadsheet, then, in essence, they have created a model to validate the results of another model. It is an algorithm testing another algorithm with no test to see if either reflects reality. Results of a model can only be validated using real world data. And finally, the V&V team did not document or state the actual results of the spreadsheet test that was performed. The results achieved and level of accuracy should be clearly documented in the V&V report so that the user can get a better understanding of how reliable the results are.

The Nuclear Propulsion Technical Logistics Model V&V report did not denote a validation test in the validation section. However, a sufficient validation test was mentioned in the Systems Verification stage. The NPTL model report stated that the model was tested using prior fiscal years' actual returns and then compared against output results within this specific years. The model was determined to achieve 98 percent validity on all tests. This was the only information provided on results validation.

The Environmental Restoration Model results were validated by using real world data from fiscal years 2000 to 2003. Actual model performance was documented representing a 14 percent difference. Every six months the model outputs are validated during data submit reviews. The level of error for the Restoration model was considered to be reasonable considering the high amount of uncertainty associated with the environmental program requirements.

c. *Recommendations*

Results validation is a critical step in the VV&A process. It is "conducted to determine the extent to which the M&S addresses the requirements for use, to identify how realistic the outputs are, and to document how well the simulation fits the intended use" (DON, 2004b, p. 18). There is no prescribed method to determine which techniques or procedures to employ to validate a particular model. However, the more objective the technique, the more confidence and trust the model users will have in the simulated results. The main purpose of results validation is to build user confidence in the performance model by demonstrating the model actually works properly. The goal is to enhance the credibility of the resource allocation tool. Therefore, there are two recommendations for this phase.

Recommendation 9. First, it is recommended that the tests used to validate the results of the model be reality-based and extensively described with the results of the test explicitly stated in the report. This stage in the V&V process is crucial for stakeholder buy-in. This is the phase where the model's performance is critically examined against real world data which it was built to simulate. Revealing the actual level of accuracy the model achieved is highly beneficial to the users and accreditation team that will evaluate the model for its intended purpose.

Recommendation 10. It is also recommended that the V&V team benchmark the model output and performance data as much as possible with similar programs in the Department of Defense or industry. Similar programs in DOD can be found using the Department of Defense Modeling and Simulation Resource Repository (DODMSMO). The metrics can be benchmarked against programs with similar purposes to ascertain the suitability of program measures of effectiveness and performance.

6. Accreditation Phase

a. *Description*

"Accreditation is the official determination by the user that the capabilities of the M&S fit the intended use and that the limitations of the M&S will not interfere in drawing the correct conclusions" (Chew and Sullivan, 2000, p. 817). The Department of the Navy's definition is "the official determination that an M&S application and its associated data are acceptable for use for a specific purpose" (DON, 2004b, p. 2). The accreditation process, unique to the Department of Defense, was implemented as the final certification that the right model was built correctly and is suitable for its intended application. It represents the independent risk reducing, credibility check that the models can simulate real world behavior with sufficient accuracy to support decision-making. DOD Instruction 5000.61 directs that "all models and simulations used to support major DOD decision-making organizations and processes…shall be accredited for that use by the DOD component sponsoring the application" (DOD, 1996, p.2)

The phase begins with an accreditation ('A') package from the model proponent. At the very minimum, the package should consist of a V&V plan, a V&V reports, model development documents, and a configuration management plan (DON, 1999). An accurate depiction of how V&V process transitions to the 'A' process in VV&A is in Figure 4.2 below.

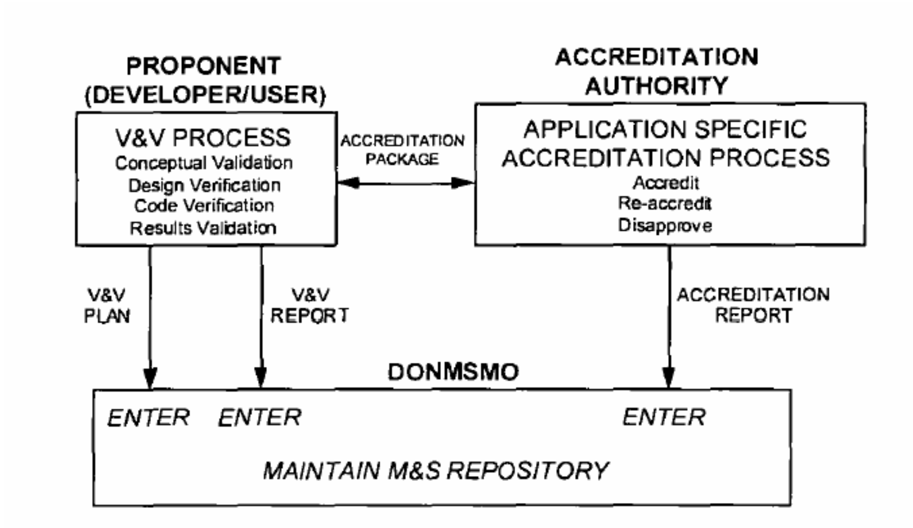

Figure 4.2 Navy M&S VV&A Process Overview (From: DON, 1999)

All of the documentation developed during the M&S V&V process becomes a part of the 'A' package that is submitted to the Accreditation Authority. Also portrayed in Figure 4.2, the V&V plan and report documentation, and eventually the 'A' decision report are supposed to be sent to NAVMSMO for information sharing purposes.

Upon receipt of an acceptable and complete package, the N81 accreditation agent commences the review process to ensure the right model was built to suit the user's purpose. To be able to make this determination, the team members must be qualified and thoroughly understand the intricacies of the performance model. The Accreditation Authority may request additional review from either a support agent, DONMSMO (or NAVMSMO), or both (DON, 1999). Additional tests may be requested, additional information may be requested, or technical assistance may be required to accomplish the certification. The overall decision as to the level of user confidence and fidelity associated with the performance model rests with the Accreditation Authority.

b. Analysis of Practice

The Department of the Navy has developed eight acceptability criteria that all performance-based pricing models must be evaluated against before a program can receive accreditation. The criteria, shown in Appendix A, consist of the following key performance parameters for the requirements models:

Performance Goals

Performance Levels

Key Drivers

Components

Design

Configuration Management

Feedback Loop

User Community

Each component in the performance-based pricing models must be linked to CNO performance goals and each program must be capable of producing cost estimates for at least four performance (or readiness) levels. For the key drivers criteria, the data, assumptions, and guidance driving the model results must be credible and certified. The data must also be documented and available for review and revision. For the component review, all major program elements, as much as possible, should be modeled and supported by a sound, written configuration management plan. Documented reasons need to be provided to justify a component remaining as level of effort. The criteria for design stipulate that the model's framework, algorithms, data sources, and assumptions must accurately reflect the validated conceptual model to produce plausible results. Also, each program needs to have a sound feedback loop to permit the validation of the model's accuracy. And finally, the model must be designed for the level of competency of the intended users and be supported by user and technical manuals.

Each criterion is assessed on a red, yellow, and green rating system to depict the model's level of sufficiency in meeting the measure. Red signifies that the evaluated area does not meet the key parameters. Green means the criterion was met and yellow expresses the need for additional effort to fully meet the user acceptability criterion. Based on the scores received for the acceptability criteria, the accreditation team recommends to the accreditation authority the decision as to whether the model should be accredited. Appendix B details the five possible decision options that can be made for accreditation. They are full accreditation (high confidence), accreditation with limitations (medium to high confidence), accreditation with modifications (medium confidence), requires additional V&V (low to medium confidence) and no accreditation (low confidence). The Aviation Depot Maintenance Model, the Nuclear Propulsion Technical Logistics Model, and the Environmental Restoration Model have all received full accreditation.

The Aviation Depot Maintenance Model was initially accredited with limitations in November 2003. The documentation provided to the accreditation team consisted of the verification and validation reports for both the ADRAM and the EDRAM, as well as supplemental inputs. Model development documentation and a formal configuration plan for these models were not included in the accreditation package. Of the eight acceptability criteria, the Aviation Depot Maintenance Model received five green ratings and three yellow ratings. The green ratings were in the categories of performance goals being linked to CNO goals, performance levels, modeled components, design, and feedback loop. The yellow ratings were in key drivers, configuration management, and user community. The reason for the yellow in the key driver area was because adequate documentation data for the data sources was not provided in the V&V report. Also the accreditation team needed to receive documentation on the pricing calculation for pricing the input. The reason for the configuration rating was that there was no configuration management plan for the ADRAM or EDRAM. And the third yellow score was because the user manual needed to be updated. Upon addressing these issues to the level of sufficiency required by the accreditation team, the Aviation Depot Maintenance Model received full accreditation in May 2004.

The Nuclear Propulsion Technical Logistics Model achieved full accreditation in August 2004. Though the performance model received a green rating in all categories, there were two weaknesses pointed out during the accreditation review. First, the changes to the force level and availability schedules change the mathematical algorithms and outputs of the model. These changes needed to be meticulously managed and documented. The second finding was that the model was "immature and the feedback mechanism must be closely monitored to establish if the data loop is correct" (DON, 2004c, p. 6). The model was fully developed in December 2003 and entered the V&V process early 2004. Therefore, it has never been used for resource allocation decisions. However, during the V&V results validation phase, the validation team tested the model using actual performance data from prior fiscal years and compared them with the NPTL model output. The only issue with granting full accreditation to the model that has achieved a 98% validity using real world data, is the fact that the accreditation rules stipulate the requirement that the model be used for a full program and budget cycle to generate execution data. The 2004 template in Appendix A, which was used to V&V the NPTL model, reads " if you have a model that was used for the first time in PR05, real world data will not be available until execution data is available at the end of FY05…full accreditation requires that Results Validation is complete" (Appendix A). The feedback loop was not really proven when accreditation was decided. Perhaps this requirement needs to be reviewed for its criticality.

In the beginning of 2004, the Environmental Restoration Model attained full accreditation status for the performance-based pricing model. The model received a green score in all acceptability categories except for Configuration Management. The yellow rating for configuration management was due to the lack of an overall, written configuration plan for the model. The plan was still under development. The model was also found by the accreditation team to have an additional weakness. Though the Restoration model achieved a green rating in the feedback loop criterion, the accreditation team cited the fact that for cost performance feedback, the cost estimating relationships had not been updated or changed since 2001. The accreditation team issued a recommendation to execute the cost modeling feedback portion of the model.

c. Recommendations

The acceptability criteria for the accreditation rating represent the key performance parameters that each model must meet to be implemented for use in the programming and budget development cycle. By definition, these are key attributes of the modeling system. In line with this thinking, the following recommendations are provided.

Recommendation 11. It is recommended that the accreditation team recommend full accreditation only to those modeled programs that meet all of the acceptability criteria with a green rating. If all of the criteria are not critical, then a priority-based accreditation system is recommended to clearly portray to program managers and V&V teams what criteria they specifically have to meet to reach full accreditation. This would benefit the program managers and V&V teams to best tailor the modeling and simulation and V&V efforts to meet the particular needs of the program within associated resource constraints. For example, if having a sound configuration management plan and a sound design were critical to receiving a high acceptability rating, then full accreditation should not be considered unless those documents are provided.

Recommendation 12. It is also recommended that the 'A' team verify and validate input data as well as user data when thorough information is not provided in the V&V documentation.

Recommendation 13. The final recommendation is for the model to have been used for at least one program and budget cycle for complete results validation to be achieved and full accreditation granted. Some stakeholders contend that "models should only be accredited once they have been used by all stakeholders in at least two budget submissions…to determine if they are useful" (Personal interview, 2005). This ensures the feedback loop is sound and a continuous improvement process is in place for performance and budgeting integration initiatives.

7. V&V TEAM ASSESSMENT

a. Description

The verification and validation team for the performance modeling and simulation project consists of a verification agent, a validation agent, and any other key individuals who have a stake in the outcome and success of the model. Though there are no definitive principles as to who needs to participate on modeling and simulation project teams, DOD and the modeling and simulation industry have increasingly emphasized the use of a systems engineering and interdisciplinary approach to "translate operational needs and requirements into a well engineered system solution" (DOD, 1998, p. 2). This is typically accomplished through the establishment of Integrated Product Teams (IPT). IPT can be defined as a "multidisciplinary group of people who are collectively responsible for delivering a defined product or process" (DOD, 1998, p.2). Developing and producing a credible performance-based pricing model suitable to inform large dollar value resource allocation decisions qualifies as a project conformable to the benefits of a multi-functional, integrated team. For the performance modeling effort, an IPT could consist of the program manager, intended users and stakeholders, and associated contractors. This approach brings about the increased possibility of generating fully integrated solutions, participant buy-in, getting it right earlier, less rework, and reduced costs and cycle time.

b. Analysis of Practice

The V&V for the Aviation Depot Maintenance Model consisted of the LMI contractor, who acted as the verification agent, the model manager from the OPNAV staff, who was the validation agent, and the model user from Naval Air Systems

Command (NAVAIR). The only representative absent from the team that should have been present was the other intended user of the performance model, FMB.

The NPTL model was also verified and validated through an integrated team. The key players consisted of the Naval Sea Systems Command (NAVSEA) model managers and users. Again, no representative from the Office of Budget (FMB) participated in the model development and V&V phase. However, this team included the members of the OPNAV N81 accreditation group. In fact, the N81 points of contacts were named as the verification and validation agents for the program performance model (DON, 2004c). This finding was of notable interest because the evaluation for accreditation is usually conducted by an independent party to assess model acceptability and accreditation recommendation (Balci et al, 2000). However, there is no known written requirement stating the accreditation agent needs to be independent of the V&V team. The issue generally addressed in the modeling community is that verification and validation should be "performed by someone other than the developer" (Arthur and Nance, 2000, p.860) to maintain objectivity and functional correctness. The DON guidance leaves plenty of flexibility in how these decisions are made. The decision for a particular program generally centers around the complexity and risks associated with the program.

Of the three models, the Environmental Restoration Model appears to have had the most thoroughly integrated team. The V&V template shows the Navy Facilities Engineering Command users as the verification and validation agents. It also reveals the other members of the V&V team. Representatives from OPNAV N45, N80, and N81 participated in the process as well as FMB users.

c. Recommendations

The analysis of the verification and validation teams developed for the performance-based modeling effort produces two recommendations.

Recommendation 14. It is recommended that FMB stakeholders are actively included as participants on the model development and V&V integrated product team. Regardless of the depth of FMB's hands-on manipulation of the model for budget formulation, attaining their buy-in and cooperation early in the process has typically been a critical success factor for government performance and budget-based initiatives (McCaffery and Jones, 2001). The models were purported to be the Department of the Navy's answer to the Budget and Performance Integration Initiative. Therefore, the budgeteers and program performance managers probably should have orchestrated the verification and validation undertaking through a coordinated effort. Often, FMB "stakeholders are unaware that a VV&A is even taking place for models in their area until the very end. And it is sometimes unclear who is in charge of the process" (Personal interview, 2005).

Recommendation 15. It is recommended, for objectivity and credibility purposes, that the verification and validation agents be different than the accreditation agent. While it may provide synergistic benefits to allow the accreditation members to participate in IPT V&V reviews and processes, the actual agents in charge of process should be independent of the team that will later certify the model for its intended use.

B. SUMMARY

This chapter has provided a comparative analysis of the OPNAV VV&A techniques practiced to certify performance-based pricing models as credible tools for resource allocation decision-making. The verification and validation activities were analyzed using a pattern-matching technique collated with the tasks typically employed by DOD and the business modeling community. 2004 guidance from the Navy Modeling and Simulation Office was also considered during the evaluation of this certification process. Based on the analysis, the following recommendations were made for future modeling efforts:

Recommendation 1.

The V&V template should incorporate Requirements V&V as its beginning VV&A phase starting with VV&A plans.

Recommendation 2.

The V&V team should link the conceptual model to the accreditation criteria.

Recommendation 3.

The activities for Data V&V should spread throughout the V&V process from requirements V&V to Results V&V per Figure 4.1.

Recommendation 4.

The Design V&V should ensure a detailed design is produced and linked to the conceptual model, user requirements, and the acceptability criteria.

Recommendation 5.

The design phase should incorporate the verification of data transformation methods applicable to the particular model.

Recommendation 6.

The V&V team should provide thorough explanations and documentation in response to template requirements instead of simply listing the applicable factors.

Recommendation 7.

The V&V team should completely describe the system verification tests and clearly present the test findings in the V&V report.

Recommendation 8.

The four data V&V activities applicable to the system verification phase should be incorporated in V&V template.

Recommendation 9.

The tests for results validation should include real world data, be thoroughly described to include adjustments and assumptions, and the test results should be explicitly stated in the report.

Recommendation 10.

For benchmarking performance data and metrics, if industry standards are not available, the V&V team should try to benchmark against similar programs in DOD using the Department of Defense Modeling and Simulation Resource Repository.

Recommendation 11.

It is recommended that only programs receiving all green ratings for the key performance parameter or acceptability criteria receive full accreditation.

Recommendation 12.

The Accreditation team should verify producer data and validate user data during the accreditation process.

Recommendation 13.

Full accreditation should be granted only after a model has successfully completed one program and budget development cycle per real world execution data.

Recommendation 14.

FMB stakeholders should participate as part of an integrated product team for model development and V&V efforts.

Recommendation 15.

V&V agents and accreditation agents should be different organizations for model objectivity and credibility purposes.

The next chapter will critically evaluate how these performance-based pricing models comply with the Budget and Performance Integration Initiative using the findings from Chapters III and IV as the basis of analysis.

THIS PAGE INTENTIONALLY LEFT BLANK

V. IMPROVING INTEGRATION

Chapter II reviewed the legislative imperatives and executive policy objectives for performance-budget integration and the Navy's decision to use performance pricing models for this purpose. Chapter III described the role and use of the performance-based pricing models in the DON's Planning, Programming, Budgeting, and Execution (PPBE) system. Chapter IV examined the rigorous verification, validation, and accreditation (VV&A) process officiated to ensure the performance models were credible and suitable for a specific application. This chapter discusses the vision of accountability expected from integrating budget and performance information through modeling and simulation and identifies potential risks to achieving that vision. Specific recommendations are provided to mitigate the causes of those risks. Finally, the chapter concludes with a comprehensive summary and suggestions for further research.

A. REVIEW OF THE GOAL

The vision of the Government Performance and Results Act (GPRA) and the President's Management Agenda (PMA) is a fully articulated and transparent performance resource allocation system that effectively delivers programs in an efficient manner. Through performance-based budgeting and linking program outcomes with dollars, the intention is to "improve the effectiveness of federal programs as measured by their actual results, and to do this by improving the performance of those programs through better management" (Mercer, 2001, p.3). The overriding goal is to promote accountability for government programs by elucidating how and if programs are achieving their objectives and the costs associated with targeted levels of performance. This is expected to reduce waste and inefficiency in government functions. The results achieved through program execution are compared with the results expected given a certain level of funding. Programs that continuously fail to achieve targeted outcomes will be evaluated for possible termination or program and resource decrements. High performing programs, those meeting specified performance objectives, will be rewarded through budget increments or simply receiving the requested fiscal resources required to execute to performance level.

The Department of the Navy has incorporated performance-based pricing models in their PPBE process, in part, to achieve the vision of GPRA and the PMA. These models, once verified, validated, and accredited to be used for resource allocation decisions, are intended to streamline the PPBE process, clearly link budgetary resources with performance information, and incentivize the overall efficient and effective management of the modeled programs.

As discussed in Chapter III, beginning with the Program Objective Memorandum (POM) for FY 2006, performance models are expected to be utilized in each phase of PPBE; from planning to execution. Assisting in the creation of Sponsor Capability Plans, the models estimate the planned requirements to perform to specified capability levels. In the programming phase, for Sponsor Program Proposals and POM development, the performance models simulate the capability levels achievable within informed constraints imposed by senior level input and resource expectations. The next phase involves the potential for the Office of Budget (FMB) to manipulate the performance model and evaluate the impacts of re-balancing, re-pricing, and adjusting execution targets (DON, 2003a).

Yet the impetus for the implementation of the revised PPBE process was to emphasize the focus on execution. It is through proper execution, tracking, evaluating, and feeding information back into the program or model that accountability can be effectively measured. MID 913 established the PPBE system and changed the programming and budget cycle to a two year process, with the off-year focusing primarily on the evaluation of program execution. During execution, program managers can use the models to develop the estimates for the next Program Review. They can also track and monitor the current progress of program achievement and instigate necessary adjustments using modeling and simulation scenarios. It also provides the opportunity to assess the performance (results) achieved of the previous year, compare it with targeted outcomes and ascertain the reason for any divergence.

Integrating performance models in the PPBE process and allocating the intended amount of effort to the execution phase through constant tracking and monitoring enhances the ability to achieve performance accountability. Employing these processes through the use of performance models makes the accountable performance more transparent and defendable. The Aviation Depot Maintenance Model, the Nuclear Propulsion Technical Logistics Model, and the Environmental Restoration Model, the three fully accredited models selected for this study, are accomplishing this, for the most part. However, there is cause for concern in the performance models' ability to fully achieve the vision of GPRA and PMA.

B. CONCLUSIONS AND RECOMMENDATIONS

Stemming from weaknesses in the integration of the various phases of the PPBE process, there are certain risks to achieving the vision. These risks take three forms: (1) the models and supporting processes such as VV&A may not be dynamic enough to deal with a transformational Navy; (2) there is an attenuated link between programming and budgeting in that the budget justification material does not accurately address the justification used for the resource allocation decision; and (3) given the rate of personnel turnover in the various offices involved in the process and the dearth of proper documentation, corporate knowledge management is threatened.

1. Dynamic Models in the Face of Transformation

Transformation is the dominant theme in the U.S. military today. It is about sustained, meaningful change in the military and organizational effectiveness in order to maintain the asymmetric advantages of the Defense Department and capably counter the known and unknown advantages of potential competitors. It involves transforming business processes and developing increased efficiencies in supporting organizational strategy within an environment of finite resources.

The Department of the Navy's Sea Enterprise program and Resource Management Concept are the primary vehicles driving the Navy's business process strategic transformation. The concept "captures efficiencies by employing lessons from the business revolution to assess organizational alignment, target areas for improvement, and prioritize investments" (DON, 2002b, p.1). Sea Enterprise expresses the need to continuously improve the efficiency and effectiveness in Navy programs. Through the use of performance-based pricing models, savings are expected to be achieved by attaining of program goals at reduced costs. By making clear the linkages between the costs to perform program activities, those activities and their resultant outputs, and then the outcomes that are affected, performance models can be employed to assess and address efficiency and effectiveness. Surpassing the targeted level of performance in a program at the same costs clearly demonstrates the markings of a well-managed program. Accountability is not just a measurement of whether a particular program has met its stated objectives, how the results are achieved and the process efficiencies associated with those results count as a significant indicator of program performance as well.

In Chapter III, it was noted that the models are used well in the planning and especially in the programming phases of PPBE. To a lesser extent they are used in budgeting, primarily to ensure that pricing decisions are properly reflected throughout the budget. Little use was recorded in the model documentation regarding the tracking of execution data and the feedback into the model parameters. This is an area of concern: the manner in which models are constructed, used, verified, validated and accredited (VVA) and the organizations responsible for those tasks have not developed a process to ensure the model is responsive to a transforming Navy. A hypothetical example will illustrate:

Most models are variations of the general linear relationship with multiple independent variables as reflected in Equation 5.1.

$$Y = a + b_1X_1 + b_2X_2 + ... + b_iX_i \quad \text{[Equation 5.1]}$$

Where Y is the dependent variable (capability, readiness level, or cost), X_i are the various independent variables and b_i are the coefficients representing the relationship between X and Y. The models are developed based upon those factors most predictive of the dependent variable using actual fleet data as described in previous chapters. The variables, data, formulation and testing undergo a strict VV&A process also described earlier. Any such mathematical formulation, however, is useful as a predictor only when the new data fall within the range of the data used to develop the model. However, concomitant to the use of the model, the underlying activities and relationships are transforming.

For example, ships operating under the Fleet Response Plan do not experience the same interdeployment training cycle the ship operations model presupposes. Likewise, the aviation depot maintenance models were not built with full consideration of the AIRSpeed initiatives that are radically streamlining aviation depot processes. So, if a model is developed assigning X_1 as depot turnaround time and X_2 as a required inventory level of ready components, a dramatic change in X_1 outside the bounds of the data used to develop the model will alter the relationship between X_1 and X_2. If the depot can turn the item around ten times faster, we obviously need less of that component since less are in the repair portion of the pipeline.

At present, the models are updated to reflect current values of X_i, but the process is not evident to periodically question the relationships between the several X_i that comprise the underlying mechanics of the model (i.e., the b*i*). Thus, as the activities transform, those models that incorporate those activities (or perhaps merely an output measure) must be attuned to those changes and respond accordingly. There are presently two impediments to model responsiveness.

a. Impediments

One impediment is the sheer complexity of the Navy and fractured assignment of responsibility for the various aspects of the processes of model development, construction, VV&A, use, and modification. For an aviation support program, for example, programming is performed within N4, the process is controlled by N8, NAVAIR owns the depots performing the maintenance activities, the fleet has the readiness requirement, and FMB owns the budgeting process.

Picture a change such as the one hypothesized above: NAVAIR's AIRSpeed[1] initiatives are making dramatic improvements in depot operations, cost and cycle time. These changes may result in values for a given X_i which are outside the

[1] NAVAIR AIRSpeed is a corporate-wide initiative to implement better business practices and tools to achieve cost efficiencies, reduce depot cycle time, and eliminate non-value added work activities.

bounds of the data used to construct the model and therefore are of questionable predictive quality. The good work of AIRSpeed may be unintentionally mitigated because resource allocation decisions are based on obsolete relationships among causal factors.

It is not discernible that developers and users assess those changes, reevaluate relationships, and revise the models under current processes. That level of inter-organizational integration was not evident in the data although it is required: these "changes to the M&S [models and simulations] must be verified and validated to determine the impact, if any, on the intended use" (DON, 2004b, p.20)

The second impediment is, ironically enough, the very process for validating, verifying and accrediting the model. The process is viewed as a gate which the model developers and stakeholders must clear in order to use the model. Once cleared, there is little incentive to return to the process. Yet, if underlying assumptions are changing due to the transformation of naval structure, doctrine or business reform, then the VV&A process should be revisited. A thorough and rigorous initial VV&A process is healthy. But it becomes a bureaucratic deterrent to changing the model.

N81 recognized this shortcoming through the results of the PR-05 and POM-06 performance modeling efforts and adjusted Program Review 2007 accordingly. To mitigate the risks associated with constant changes in the variables and underlying assumptions to the model, the Navy has implemented guidance to force the fully accredited programs to undergo routine re-accreditation every three years. This entails a complete re-examination of the model by the model stakeholders and the accreditation team. Additionally, the guidance mandates that specific processes be developed and revealed to the accreditation team detailing the "mechanisms" in place to ensure the model data remain accurate within the constructed interval range. If processes change significantly within the three year time-frame, the model managers are expected to commit the model for an "emergent re-accreditation." For the emergent review, only the aspects of the model that are affected need to be re-verified, re-validated, and re-accredited. Only time will tell how well model managers comply with the guidance.

b. Recommendations

These changes to model guidance show that the Navy is committed to the continuous improvement of the performance-based modeling and VV&A processes. An additional recommendation to ensure compliance is to have the stakeholders review these aspects of the model more frequently, particularly in the odd-numbered PPBE years when the focus is on execution because it is during the execution review that these changes will be noted.

Another potential remedy or factor that will enhance the Navy's efforts to control these risks lies in the model configuration documentation and those responsible for it. It is recommended that full accreditation should not be granted to programs without an effective, thorough, and complete configuration management plan. Configuration management is the critical element that ensures the continued integrity of the model and its simulated output. Through an effective configuration management plan, program managers can manage and document the functional and physical aspects of the model and track and record any changes to modeled configuration items (CI). The documentation identifies the various assumptions and technical data associated with requirements and design features and details the procedures for communicating and incorporating CI changes. Incomplete documentation, with un-delineated assumptions, relationships, and version specifications impedes the VV&A process and can result in an unnecessarily expensive and time consuming evolution. However, developing and maintaining an up-to-date, comprehensive configuration management plan through accurate and complete documentation endues the increased potential for program managers to readily modify model configurations and facilitates the re-verification and re-validation effort.

Augustine's Law states "No change is a small change" when it comes to complex systems. Therefore, ensuring that a sound configuration plan exists is critical to the credible performance of these models. Recommendations were provided in Chapter IV to ensure that models do not receive full accreditation without achieving a green in each of the key performance parameters composing the acceptability criteria. This implies that without a complete and thoroughly documented configuration management plan, incorporating an updated list of model assumptions, performance models cannot be granted full accreditation.

A third recommendation is to incorporate the review of the configuration documentation and its underlying assumptions during the yearly MPVT event. It is also recommended that the various data source providers attend the review session to ensure consensus and continued understanding of model component relationships and its simulation capabilities. If these recommendations can be incorporated, then the VV&A repeat can be an easier, quicker process. Also, depending on the complexity of the change, it may be beneficial to perform the VV&A in this forum while all of the key stakeholders are present to ensure adequacy and conformity. The crucial building block in successfully improving the model's performance in the environment of transformation is thorough, complete, up-to-date, and accurate configuration documentation.

2. Corporate Knowledge Management

Another potential risk to achieving the GPRA and PMA vision emanates from the high rate of personnel turnover inherent in military organizations. Military personnel rotate from their jobs typically every two to three years. Quite commonly, the knowledgeable participants involved in the model development and verification and validation process are not the current managers of the performance model. This coupled with the dearth of proper documentation of M&S development efforts, VV&A procedures, and most notably the configuration management plan mentioned in the previous section, threatens the corporate knowledge in the management of performance-based models.

Fully understanding how the model is designed and how the key model drivers interact and impact the targeted outcome is critical in properly managing the performance of the program. Detailed knowledge of data sources and the procedures employed in generating model input data is among the most significant corporate essentialities in program administration. Without this information, program managers may not know the right questions to ask when analyzing the accountability aspects of the outcomes achieved. As Chapter IV addressed the issue, these were the areas most notably deficient in documentation. VV&A documentation commonly listed the data sources but did not describe or expound on the V&V of the data and how the data sources derived the information.

As mentioned in the previous chapters, the Navy Modeling and Simulation Management Office (NAVMSMO) serves as a central repository for all Navy modeling and simulation documentation. An abbreviated list of NAVMSMO's responsibilities is as follows:

- Guide the efforts within the Navy to achieve the Navy's M&S vision
- Maintain a centralized M&S information distribution system and catalog of Navy M&S resources
- Establish policy; develop and/or maintain the instructions and/or standards necessary to manage the Navy's M&S program
- Review resource on new model development
- Identify, investigate and document M&S requirements, including emerging M&S and information technologies, and assist in prioritizing Navy-wide programs (NAVMSMO, 2005).

The Navy is currently required to submit model development and VV&A documentation to NAVMSMO for storage. However, there is a weak link between NAVMSMO and the Navy's performance-based pricing model program. Certain individuals interviewed for this research presumed NAVMSMO's role applied only to the Navy's warfare systems modeling and simulation efforts. Yet, SECNAVINST 5200.40 relates the requirement to all DON models, including all "M&S which are used significantly in supporting the development of either the DON Program Objectives Memorandum (POM) or Analysis of Alternatives" (DON, 1999, p. 3). Therefore, it is recommended that the performance modeling documentation be forwarded to NAVMSMO to aid future modeling efforts and to assist in the sustenance of corporate knowledge.

3. Performance-Based Programming versus Performance-Based Budgeting

The goal of GPRA and the PMA's Budget and Performance Integration Initiative is to make government programs more accountable through transparency in performance-based budgets. A performance-based budget is "an integrated annual performance plan and annual budget that shows the relationship between program funding levels and expected results" (Mercer, 2002, p. 2). As communicated in the results-oriented framework (introduced in Chapter II) in Figure 5.1, it is a functionally structured budget linking performance, cost, and resource information to improve the ability for program managers to achieve targeted objectives.

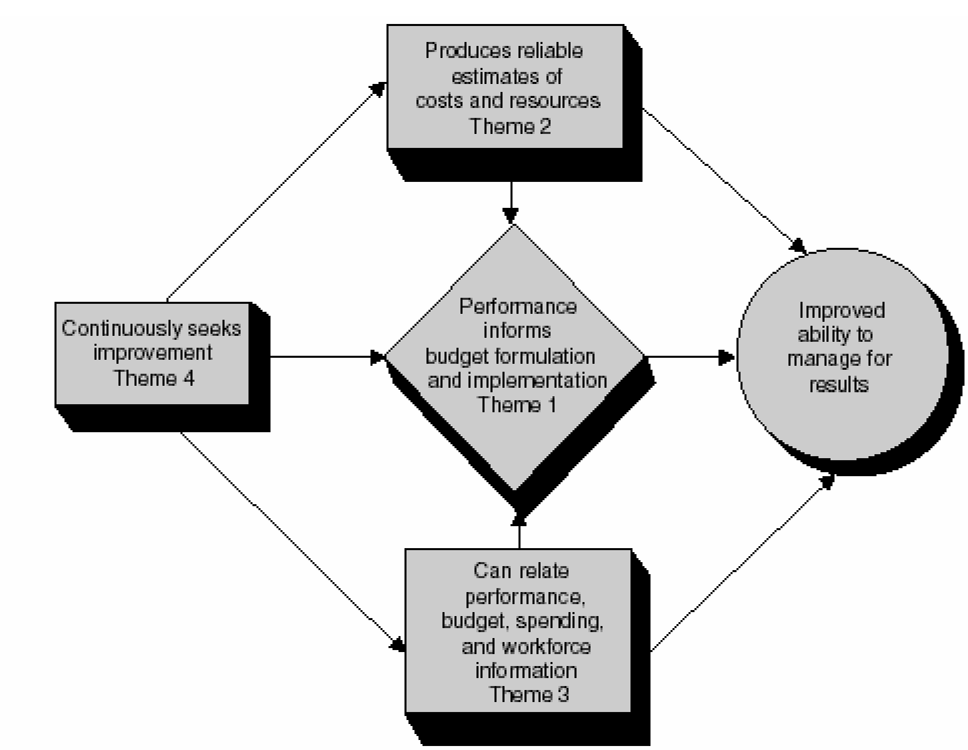

Figure 5.1 Results-based Budgeting Framework

Performance-based pricing models are one of the primary tools implemented in the Navy to meet the expectations of performance budgeting and the President's Budget and Performance Integration (BPI) Initiative. The models link performance and cost information with budgetary resources and provide a defensible strategy to align program results with organizational goals. The Navy has taken significant steps in satisfying the requirements of the BPI and performance-based budgeting approach. However, key features of the results budgeting construct have not been fully implemented and occlude the achievement of the GPRA and PMA's vision of accountable performance through budgeting.

As portrayed in the performance-based budgeting framework (Figure 5.1), program performance is enhanced through the meticulous management of the four supporting themes. The arrows indicate the facilitation of theme one through the continued accomplishment of themes two, three, and four and the persistent focus on improving the capability of all themes to achieve better program results.

a. Theme One

Theme one, 'Performance informs budget formulation and implementation,' represents the main objectives of the BPI Initiative. The budget should be linked with performance information. In other words, program performance information should advise resource allocation decisions during the development and execution of the budget. In the formulation phase, program managers should be able to determine the capability or readiness level achievable given resource and political restraints. Based on certain funding levels, managers need to be able to articulate program impact using credible and defensible tools. The Department of the Navy's performance models answer this requirement through the simulation of various resource constrained capability levels and what-if scenarios.

Budget implementation concentrates mainly on tracking and monitoring program performance during budget execution. In this phase, the manager assesses program achievement within the triple constraints of cost, schedule, and performance objectives. In analyzing whether the program is on track with spending the predicted amount of resources in relation to the scheduled and predicted performance level, managers can take necessary measures to realign the program to attain the expected outcome. In the execution phase of PPBE, program managers can review three functional areas. They can perform gap analysis on the results achieved in the previous year with targeted program goals. They can focus on monitoring the execution of the program in the current year, as discussed in theme one. And they can concentrate on assessing and re-evaluating the programmed and budgeted resources for the next Program Review (PR) Cycle.

Interviews conducted during this research revealed that the majority of a program manager's time was spent in planning and programming. Even though the MID-913 reform of the PPBE process emphasizes budget execution, a few managers felt that they had too little time to devote to monitoring and tracking the execution of their programs (Personal interview, 2005b). They expressed that they were constantly in the P, P, and the B phases of planning, programming, budgeting and execution. The on-year demands a great deal of focus on POM development and the off-year seems to draw more emphasis to the upcoming PR. Also, in the search for specific DON guidance for budget execution, the only information obtained provided direction on the preparation and submission of program and budget estimates for the PR cycle. The guidance stressed the use of performance-based pricing models to identify significant issues and link performance with the budget to achieve program accountability (DON, 2005).

The concern with spending too little time on the execution process and focusing primarily on planning and programming is that there is a lack of opportunity to improve program and model performance. The underlying goal of GPRA and the PMA is to improve the performance and accountability of government programs. The models and budgeting initiatives currently in place are being used to satisfy the requirement of linking dollars with expected results. However, there is limited evidence that enough attention is dedicated to ascertaining if the right amount of dollars were linked to the right amount of resources at the right costs.

b. Themes Two and Three

Themes two and three are the critical underpinnings of the BPI theme. Theme two "Produces Reliable Estimates of Costs and Resources' and theme three "Can relate performance, budget, spending, and workforce information' address the quality, reliability, and accuracy of the information to support resource allocation decisions. These themes define the core criteria for achieving a true performance-based budget. One important feature is the need to produce useful cost estimates through a valid cost allocating methodology. Preferably, the use of activity based costing, or some method of developing reliable unit costs, is recommended to inform performance trade off decisions and conduct strategic cost analysis. "If the Department intends to transform, it must look forward (using managerial accounts) in building budgets, not backward (using financial accounts) or at what worked last time (using budgetary accounts)" (Candreva, 2004, p.9). With a good understanding of the unit costs of program requirements, managers can more effectively assess costs versus benefits and determine affordability of desired requirements.

Based on the analysis of the three fully accredited performance-based pricing models, the Navy appears to use valid cost estimating methodologies to price the readiness levels of their programs. The Aviation Depot Maintenance Model is based on working capital fund procedures that set rates based on unit cost. The NPTL model uses unit costs for man-years and the Environmental model incorporates a parametric cost estimating technique based on unit costs and parameter relationships. Though this assessment is based on the analysis of only three out of over forty performance models, these are the only models to achieve full accreditation to date. This inspires the positive conclusion that producing accurate and reliable costing data is of significant importance in a model receiving a high credibility rating.

(1) Risks in Performance-based Budgeting. The major risk in performance-based budgeting derives from the other features captured in themes two and three. Performance budgets require the development of performance measures and indicators to enable managers to determine and report program performance. They are also premised on a more functionally-based budgetary construct. Instead of allocating resources to cost objects, dollars are supposed to follow program activities or tasks that need to be performed to achieve the desired outcomes. A performance budget reveals the linkage between the costs of these activities and program outputs and outcomes. Additionally, these performance metrics are supposed to be reflected in the budget justification material (DOD, 2002b). The performance measures employed in the performance-based pricing models should be displayed in the congressional budget documents. The Department of Defense specifies that "in the budget review, the USD

(C) will use the metrics that components submit as part of the budget estimate submission to make informed resource allocation decisions" (DOD, 2003, p.7). The budget request is justified to Congress based on the performance measures articulated in the budget material.

The analysis of the three modeled programs, reveals significant improvement needs to occur in this area. The attenuated link between programming and budgeting mentioned previously manifests from the paucity of complete performance measures in the budget justification material when compared with the detailed measures entwined in the program's performance model. In short, the measures justifying the budget request are not often the measures used to arrive at the budgeted amount.

(2) Performance Measurement Concerns. There are two concerns with the Navy's use of performance measures associated with budget justification and resource allocation decisions. The first is that not all modeled programs have included performance measures in the budget justification material. The Nuclear Propulsion Technical Logistics (NPTL) Model appears as just one line item in budget activity 1B5B with total estimated dollars allocated for the year. There is no indication of how resources will be spent to meet a targeted performance level or outcome. There is no indication of critical activities associated with the performance and how these activities drive the costs. Therefore, it would seem evident that OSD (C) is not incorporating program performance measures in their decisions to allocate resources.

Another concern is quality of performance measures that exists in the budget documents for the modeled programs. The metrics listed in the budget tend to be limited in scope. All of the reliable and relevant performance measures incorporated in the model to deliver a certain capability level are not fully integrated in the budget justification material. The only measures identified (if any are identified) are the output measures of the model. For example, for the Aviation Depot Maintenance model, the performance measure in the budget is the number of airframe or engine inductions. This is supposed to provide FMB with enough information to make resource allocation tradeoff decisions and to ascertain whether the programs are achieving its goals of fleet readiness. To create an effective accountability system, decision-makers need a healthy balance of performance indicators that measure results and processes. "A balanced family of measures can evolve into a powerful system for executing strategy" (Epstein and Birchard, 2000, p. 145).

(3) Process versus Results. There are many cost and performance drivers in the model that interact and impact the output and the outcome of the program. Focusing only on the results measure stifles the ability to determine accountable performance and manage the program more efficiently and effectively. It also reduces the quality of performance information used to develop budget estimates. As previously discussed, the goal of incorporating these performance models in the PPBE system is to develop the capability to achieve productive efficiencies and cost efficiencies. It is to achieve productive effectiveness and cost effectiveness. It is to inform resource allocation decisions by providing timely, accurate, and complete activity-costed information to show how dollars will link to program results. Making decisions based solely on the output or results of a program is much like trying to derive information from the scoreboard in a baseball game. As Shank contends, "the scoreboard tells a player whether he is winning or losing the game, but tells him little of what he is doing right or wrong in the mechanics of baseball" (Shank, 1993, p. 138). Strategic decision-making requires a balance of performance activity measures and performance results measures. In the Aviation Depot Maintenance example, if the program achieves the targeted outcome, it is important to know what critical activities were involved to bring about that outcome. Through this information, resource allocation decisions can be enhanced, as well as the overall management of the modeled program.

c. Theme Four

The final theme, "Continuously Seeks Improvement' is framed around improving the performance of themes one, two, and three in order to better manage for results. The theme deals specifically with finding new and better ways of achieving program goals and of achieving cost efficiencies. The basic premise is that "unit cost measures can help managers see tradeoffs between competing needs" (GAO, 2001, p. 13) when combined with effectiveness measures. The underlying focus is to continuously strive to improve process efficiencies, cost efficiencies, and program effectiveness. The preceding sections discussed the current challenges to continuous improvement of program and of the performance-based models. Perhaps it is sufficient to conclude with the sentiments of an interviewed stakeholder, "there is nothing forcing the models to improve" (Personal interview, 2005).

C. LIMITATIONS OF STUDY

Every study has its limitations and this is no exception. This study examined only three of the forty plus performance models developed in the Department of the Navy. The author selected those models that had successfully completed the VV&A process and achieved full accreditation. While these are the only examples of fully accredited (high confidence) models in the DON, they may not be representative of the remaining models that have successfully completed the VV&A process. The recommendations here are based on the sample and some recommendations may or may not apply to the other performance-based pricing models being used for resource allocation decisions.

This study also considered mostly archival data. Those data were augmented with data from interviews with people involved in the processes of model development, VV&A, and use. Ideally, these data would be augmented with more interviews and observational data as the model is used in the various phases of PPBE.

Finally, both the PPBE process and the use of models in support of it are still in their infancy-or at best, toddler stage. These fully accredited models, to some extent, broke new ground and the processes were refined concurrent with their development. Thus, the remaining models will certainly follow a different path to accreditation. This may render some conclusions moot, but conversely increases the urgency with which other recommendations should be considered and implemented.

D. RECOMMENDATIONS FOR FURTHER RESEARCH

The author has identified five areas worthy of further research. First, the verification, validation, and accreditation (VV&A) process of the remaining performance-based pricing models that have received accreditation could be analyzed. Findings and recommendations could be compared with the recommendations provided in this study. Also, critical success factors and/or process impediments could be identified to aid future performance modeling efforts.

Second, the accreditation process could be thoroughly reviewed. Accreditation is a DOD-specific term and process. More research to assess the completeness of the acceptability criteria and the manner in which the models are evaluated would be a benefit to the Navy. The researcher could evaluate the possibility or benefit that may or may not arise from a weighting system and how the V&V effort could better tailor their processes.

Third, the basic building blocks for performance-based pricing models and performance-based budgeting are the performance measures consociated with the program. A more thorough analysis could be conducted to assess the existing performance metrics integrated with the performance models. Recommendations of more relevant and executional performance measures could be provided to improve the models and the information used for resource allocation decisions.

Fourth, the costing methodology of the three fully accredited models was briefly touched upon to identify the procedure. A more thorough analysis of the various costing methods employed by the various performance models with recommendations for possible improvement would benefit the performance modeling effort.

Research in any of these four areas would be an extremely valuable contribution to the Navy's attainment of Budget and Performance Integration Initiatives.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX A.

Performance/Pricing Models
Verification & Validation (V&V) Template

Version 2.0, February 2004

This template will be completed for all models and submitted to the Performance/Pricing Model Accreditation authority (OPNAV N8). Attach all documentation that supports your Verification and Validation (V&V) effort. More details can be found in SECNAVINST 5200.40, VV&A of Models and Simulations.

In the context of programming and budgeting, the purpose of conducting a VV&A is to establish confidence or trust in the model or methodology being used to generate requirements. The V&V Template is a tool to collect the evidence necessary to establish the credibility of the model for its specified use.

Date of completion for this report: ______________________
Responsible author: ______________________
Author's organization: ______________________

Performance/Pricing Models
Verification & Validation (V&V) Template

Version 2.0, February 2004

1. Model Identification

Model name: ____
[Note: this name is automatically placed in each page footer when the document is printed.]

Version or release:____

Responsible verification agent:__________________________________
Agent's organization: ____
(Identification of the individual responsible for managing the verification effort and compiling the results)
Responsible validation agent:_____________________________________
Agent's organization: ____
(Identification of the individual responsible for managing the validation effort and compiling the results)

V&V Team Information (name, organization, phone and email):
- Proponents/owners:
- Users:
- Independent agent (if applicable):

Name	Organization	Phone	Fax	E-mail

2. Model Description and Background

```
Fully understanding the Modeling and Simulation (M&S) development requirements
is essential for the VV&A effort.  These requirements define the functionality
and capability, which the user requires of the model or simulation system.
They also serve as the foundation against which the simulation will be verified
and validated.
```

Please identify the acronyms used in describing the model anywhere within this completed template.

ACRONYM	EXPLANATION

A. Briefly describe the model or simulation and the program(s) the model supports.

B. This a new model, legacy model (detail the extent of the VV&A actually performed, or indicate "model used for x years with little or partial VV&A"), a model still under development, or a change to an existing model?

C. What is the history behind development of the model?

D. Summarize aspects of past V&V and/or past M&S that may impact accreditation. Provide a copy of any VV&A documentation. If the model has been formally accredited or otherwise formally approved for a specific application or set of applications, provide the documentation demonstrating formal accreditation or approval.

E. Who uses the model? Is the model designed and developed for the level of competency of the user for its intended purpose? Are there supporting documents such as user's manual, technical manual, and/or reference guide? Please either attach copies to this template or provide references.

F. Describe the model's linkage back to approved CNO goals. Typical references include sections of documents such as: Strategic Planning Guidance, Fleet Manning Documents, DOD Instructions, etc. If CNO goals have not yet been established, state so and provide any plans in place to create them.

G. Define the model's performance levels, components of the performance levels, and describe how they were developed. Demonstrate how the model is linked to readiness or other performance metrics. Ideally, performance models should have at least four performance level options. If your model does not have at least four performance levels as a result of an issue specific to your program, see your N81 Model Representative before continuing with this V&V.

H. Additional comments:

3. Conceptual Validation

The conceptual model serves as a bridge between the defined requirements and the M&S design, providing the developer's interpretation of the requirements to which the model or simulation will be built. The conceptual model is a statement of assumptions, algorithms, and architecture that relates the elements of the model to one another for the intended applications of the models or simulations.

A. Was a conceptual model developed prior to developing your model? Provide a graphic representation of the model with a written description explaining the process.

B. Drivers: List and describe the model drivers. Examples of model drivers include -- but are not limited to -- assumptions, OSD/USN/USMC policies and guidance, and output from other models.

1) Identify the key drivers below and describe why they are significant. Complete a Key Drivers V&V page for these key drivers (attached at the end).

2) Are the assumptions, policies, or guidance represented by input variables or are they fixed? Who determines whether to accept or changes these drivers?

3) If output from other models is used, provide VV&A or other documentation that validates the use of this input.

C. Components: What budget categories do you consider to be separate model components? Model components represent categories and sub-categories for which separate cost estimates are produced. Examples of model components include: Personnel (direct and indirect), projects/contracts, materials, equipment, maintenance, etc. If a POA&M has been generated addressing the future modeling of components, please attach a copy to this template.

1) List and describe the model components. Include a short name for each component and use the short name in column one of the table in paragraph (2) below:

2) Provide your program's total obligation authority in tabular format (see example table below).
In the first column list the model components and any associated subcomponents. List all subcomponents as a separate row entry.
In the second column, list the dollar amount allocated to each respective component and subcomponent.
In the third column list the percentage of the component that is explicitly modeled (in terms of the portion of the total dollar amount attributed to that component). For example, if the personnel component of a program is allocated $100 million (see Component 1 in table below) and $20 million of that amount is modeled using a cost estimation model with the remaining $80 million estimate based on a level of effort approach, then the column three entry would be 20%.
In the fourth column briefly characterize the modeling approach, or lack thereof, used for the component. If a component is not modeled, explain the reason and any steps in place to model those components. Include supplementary notes for clarifications or amplifying information if needed.

Component (short name)	**Total Cost Estimate**	**Percentage of the Component modeled**	**Modeling Method**
Component 1	$100M	20%	Cost estimating relationships
Component 2	$240M	75%	(as below)
Subcomponent 2.1	($120M)	50%	Price History/Analogy
Subcomponent 2.2	($120M)	100%	Engineering Estimate
Component 3	$ 60M	0%	Not modeled – Level of Effort
TOTALS	**$400M**	**50% - weighted average of Total Cost Estimate**	

Component (short name)	**Total Cost Estimate**	**Percentage of the Component modeled**	**Modeling Method**
TOTALS			

D. Outputs: What does the model actually produce?

1) List and describe the model outputs:

2) Demonstrate how the model outputs provide information relevant to resource allocations. Does the model have the ability to determine the requirement and price beyond the programmatic level down to the claimant/activity level?

E. Have metrics been developed to benchmark performance and pricing against industry standards or other accepted standards?

F. Describe any additional steps taken (not included in the above) to validate the conceptual model.

G. Outcome (only required for models currently under development):
Describe corrective actions planned if results not satisfactory, leading to a repeat of this validation for a later model version; or, indicate a determination that results are satisfactory; or, document any modeling workarounds planned that will avoid or minimize impacts from unsatisfactory results at this stage and will allow the work to proceed.

H. Additional Comments:

4. Design Verification

The M&S functional design is verified against the conceptual model to ensure that it accurately reflects the validated concept and associated requirements.

A. Was design verification done during the model development process? Provide a graphical representation of the model's design with a description. Identify the source of the diagram (e.g., derived from original source material, developed as part of the V&V process, etc.).

B. Demonstrate how the design meets the purpose/objectives/requirements developed in the conceptual phase.

C. Describe any additional steps taken (not included in the above) to verify the model's design.

D. Outcome (only required for models currently under development).
Describe corrective actions planned if results not satisfactory, leading to a repeat of this validation for a later model version; or, indicate a determination that results are satisfactory; or, document any modeling workarounds planned that will avoid or minimize impacts from unsatisfactory results at this stage and will allow the work to proceed.

E. Extent of Previous V&V:
If a V&V process has been performed, detail the scope of the V&V performed to date. This includes the portion or percentage of the model which has been reviewed or examined as part of the formal V&V process. The following list illustrates examples of ways to itemize the scope of a typical V&V process:

____% of code reviewed/subjected to static test methods, etc.
____% of models/functions/etc. demonstrated to perform as expected.
____% of boundary condition inputs examined
____% of input range examined in results validation
____% of inputs for which credible/authoritative data sources were identified

5. System Verification

System verification is the formal (i.e., documented) test/review process by the M&S proponent responsible for determining that the M&S accurately represents the functional design and has traceability to the conceptual model and the system requirements.

A. Model Design: Explain the model's design.

1) Provide a high-level diagram of the model as used, depicting inputs, outputs, process elements, performance feedback loop(s), and cost feedback loop(s).

2) Describe the process of how the model works, referring to the diagram produced above. Ensure program-specific terms and acronyms used are included in the glossary in part 2 above. Provide in the discussion any important assumptions and key algorithms used by the model. Ensure the following elements are addressed in the description:

How are performance and pricing determined during the programming phase and how do these two elements of the model interact?

How does actual execution data, both pricing and performance, feed back into the model? How is the model changed to reflect this data?

Demonstrate how the model is linked to readiness or other performance metrics.

Where and how are the results of the model incorporated into the Program/Budget Information System (PBIS)?

What policies influence or constrain the model's design?

How are the various model processes depicted in the design diagram implemented in the working model? (i.e., electronic spreadsheets, web-based data entry and collection, manual data calls, Java code, .net architecture)

Is the output artificially constrained by budget/financial considerations?

B. What test procedure is used to demonstrate model compliance to requirements? Provide documentation/test results.

C. How susceptible is the output to fluctuations across models/tools within and outside the system?

D. Describe any additional steps taken (not included in the above) to validate the conceptual model.

E. Outcome.
Describe corrective actions planned if results not satisfactory, leading to a repeat of this validation for a later model version; or, indicate a determination that results are satisfactory; or, document any modeling workarounds planned that will avoid or minimize impacts from unsatisfactory results at this stage and will allow the work to proceed.

F. Additional Comments:

6. Results Validation

Results validation by the M&S proponent/owner is the formal (i.e., documented) test/review process that compares the responses of the M&S with known or expected behavior from the subject it represents, in order to ascertain that the M&S responses are sufficiently accurate for intended uses.
This step can only be completed if real world data is available. For instance, if you have a model that was used for the first time in PR05, real world data will not be available until execution data is available at the end of FY05. Full accreditation requires that Results Validation is complete.

A. Provide documentation comparing the model's actual results to the expected results.

B. What errors were found and how were they corrected?

C. Describe any additional steps taken (not included in the above) to validate the conceptual model.

D. Summarize conclusions reached. Describe corrective actions planned if results not satisfactory, leading to a repeat of this validation for a later model version; or, indicate a determination that results are satisfactory; or, document any modeling workarounds planned that will avoid or minimize impacts from unsatisfactory results at this stage and will allow the work to proceed.

NOTE: If this step cannot be completed because real-world data is not available, describe the actions that will be taken to complete. What actions will be taken to incorporate any changes between System and Results validation (fund migration, unexpected events during execution, changes in performance goals, etc.)?

E. External Feedback (for ongoing validation): Demonstrate how the Navy's accounting system does/does not allow for the ability to track amount programmed vs. amount budgeted vs. amount executed for your model. What changes are required to the accounting system if it does not support your model from programming through execution?

F. Subject Matter Expert (SME) Involvement: If SME assessments were substantially used as the basis for model or data correctness or acceptability, identify the SME(s) and document their credentials below.

G. Additional Comments:

7. Configuration Management Plan and Model Management

System Configuration Management is the process through which model upgrades, changes, and maintenance are recorded, communicated and controlled. A written Configuration Management Plan is required for full accreditation.

A. Is there a written Configuration Management Plan that addresses the following questions?
If so, please attach; if not, when will the written plan be complete?

B. Describe the process for suggesting, adjudicating, and prioritizing changes to the model. Discuss the methodologies to ensure changes are documented, tracked and version control is observed?

1) Who approves changes to the model?

2) Is there a method to report status of these changes to those who have an interest?

3) Are there requirements management to ensure improvements and modifications are made according to the priority of the requirements?

4) What resources exist for life cycle support of the model? Have resources been identified and allocated?

C. How will the model's output be subject to periodic reviews and evaluation?

D. List and describe and additional Configuration Management or Model Management elements.

8. Accreditation Report Evaluation Summary

The information provided in the preceding sections forms the basis for the accreditation recommendation. This section lists the evaluation categories and evaluation criteria against which the model will be rated. As such, this section provides an opportunity to include amplifying information, not previously captured, which may affect the ratings assigned.

A. Performance Goals

Criteria	Rating Scale
For each program, modeled components are linked to CNO Performance Goals.	GREEN: Linked to CNO goals
	YELLOW: CNO goals not yet established
	RED: Not linked to CNO goals

Provide below any information not previously provided that may affect the accreditation rating assigned for performance goals.

B. Performance Levels

Criteria	Rating Scale
For each program, the model produces costs for at least four performance levels.	GREEN: Model has four or more performance levels
	YELLOW: Model has 2 or 3 performance levels
	RED: Model produces only the full cost

Provide below any information not previously provided that may affect the accreditation rating assigned for the model's performance levels.

C. Key Drivers

Criteria	Rating Scale
For each program, key drivers (data, assumptions, and guidance) are credible and subject to review and revision.	GREEN: All data is valid or certified
	YELLOW: Most data traceable to certified source; data reviewed
	RED: Key drivers are arbitrary or best guess, data not reviewed

Provide below any information not previously provided that may affect the accreditation rating assigned for the model's key drivers.

D. Components

Criteria	Rating Scale
For each program, as practicable, all components are modeled.	GREEN: As practicable, all components are modeled
	YELLOW: As practicable, a POA&M is in place to model all LOE components
	RED: No plan exists to ensure all LOE functions are modeled

Provide below any information not previously provided that may affect the accreditation rating assigned for the model components.

E. Design

Criteria	Rating Scale
For each program, the model's design (framework, algorithms, data sources and assumptions) accurately reflects the validated concept to produce credible results.	GREEN: The model's design is sound and produces credible results
	YELLOW: The model's design requires some improvements to improve results credibility
	RED: The model's flawed design produces results that are not credible

Provide below any information not previously provided that may affect the accreditation rating assigned for the model's design.

F. Configuration Management

Criteria	Rating Scale
For each program, modeled components are supported by a sound written Configuration Management (CM) Plan.	GREEN: CM process for all changes
	YELLOW: Some CM processes for all major upgrades/code changes
	RED: No formal CM process

Provide below any information not previously provided that may affect the accreditation rating assigned for the model's configuration management.

G. Feedback Loop

Criteria	Rating Scale
For each program, a sound feedback mechanism exists to allow for validating the model's accuracy.	GREEN: Comprehensive feedback mechanism in place
	YELLOW: Partial feedback mechanisms in place
	RED: No feedback mechanism in place

Provide below any information not previously provided that may affect the accreditation rating assigned for the model's feedback loop.

H. User Community

Criteria	Rating Scale
For each program, the model is designed and developed for the level of competency for its intended purpose. The model is supported by documents such as user's manual, technical manual, and/or reference guide.	GREEN: User community has the ability and tools to fully utilize the model
	YELLOW: User community has some of the tools and knowledge to use the model
	RED: User community lacks adequate tools and knowledge to use the model

Provide below any information not previously provided that may affect the accreditation rating assigned for user community.

I. Supplemental Information: Attach other supporting documentation that may facilitate the accreditation process. For example, glossary of terms, model design standards, V&V standards, etc.

KEY DRIVER

VERIFICATION and VALIDATION

(Note: Complete a separate form for each key driver.
Copy this page to the end of this template as needed.)

The data examination must consider both correctness of the data and its interpretation/translation into M&S parameters.

Date V&V completed: ________________________
Responsible author: ________________________
Author's organization: ________________________

Performance/Pricing Models
Verification & Validation (V&V) Template

Version 2.0, February 2004

1. Key Driver (Data Source or Guidance) Identification

Key Driver name:

____

Version or release:

____

Originating organization:

____

Point of contact:

____

2. Basis for Confidence in the Data Source or other Document

A. Briefly describe the Key Driver (data source or document) and how it is used.

B. Who owns and maintains this source? What drives their review and update schedule?

C. Explain why the data source or guidance/instruction document is believed credible (i.e., What makes the data or the guidance authoritative?). Attach any required documentation.

D. How is the data collected and then tied to the model? Include any data transformations of units/coordinate systems, etc. for data to be appropriate for use as model input.

E. What are the known limitations and restrictions in the data source?

F. Is there an evaluation method to ensure data source or other guidance is accurate and correct? What is the frequency of any evaluations?
In addition to detecting any substantive errors, such evaluations would typically serve to identify, and correct or eliminate, typographical errors and other data corruptions, unusual data items, etc.

G. Describe any weaknesses in the data source or document and how they may influence the outcome of the model. What is your plan to implement corrections to improve credibility?

H. What is your overall conclusion as to the suitability of data set or report for use with this model?

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX B.

Accreditation Decision Options

There are several accreditation decision options available. These options are outlined in the table below along with their respective qualifiers.

	Decision	Qualifiers
	Full Accreditation (High Confidence)	Model produces results that accurately reflect requirements that are linked to performance with high confidence in model methodology.
	Accreditation with Limitations (Medium-High Confidence)	Model results are generally valid and it is sufficiently credible to support the application. Minor constraints may be placed on the model
	Accreditation with Modifications (Medium Confidence)	Constraints are placed on how the model can be used to support the application Model capabilities are insufficient and modifications are necessary. A subsequent V&V is needed to correct the deficiencies.
	Requires Additional V&V (Additional Information is Needed) (Medium-Low Confidence)	Information obtained about the model is insufficient to support an accreditation decision; supplemental V&V and/or testing should be conducted to provide the necessary information before the accreditation decision is made.
	No Accreditation (Low Confidence)	Results of the assessment show that the simulation is not fit to support the application.

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

Arthur, J. D., & Nance, R. E. (2000). Verification, and Validation Without Independence: A Recipe For Failure. *Proceedings of the 2000 Winter Simulation Conference*. [CD-ROM] (OmniPress).

Balci, O., Carr, J. T., Ormsby, W. F., & Saadi, S. D. (2000). Planning for Verification, Validation, and Accreditation of Modeling and Simulation Applications. *Proceedings of the 2000 Winter Simulation Conference*. [CD-ROM] (OmniPress).

Candreva, P. J. (Fall 2004). Accounting for Transformation. *Armed Forces Comptroller,* 7-13.

Carson, J. S. (2002). Model Verification and Validation. *Proceedings of the 2002 Winter Simulation Conference*. [CD-ROM] OmniPress).

Caughlin, D. (2000). An Integrated Approach to Verification, Validation, and Accreditation of Models and Simulations. *Proceedings of the 2000 Winter Simulation Conference*. [CD-ROM] (OmniPress).

Chew, J. & Sullivan, C. (2000). Verification, Validation, and Accreditation in the Life Cycle of Models and Simulations. *Proceedings of the 2000 Winter Simulation Conference*. [CD –ROM] (OmniPress).

Department of Defense. (1996). DOD Modeling and Simulation (M&S) Verification, Validation, and Accreditation (VV&A). Department of Defense Instruction no. 5000.61. Washington, DC: U.S. Government Printing Office.

Department of Defense. (2002a). *Management Information Decision 901: Establishing Performance Outcomes and Tracking Performance Results for the Department of Defense*. Washington, DC: U.S. Government Printing Office.

Department of Defense. (2002b). *Management Information Decision 910: Budget and Performance Integration Initiative*. Washington, DC: U.S. Government Printing Office.

Department of Defense. (2003). *Management Information Decision 913: Implementation of a 2-Yr Planning, Programming, Budgeting, and Execution Process*. Washington, DC: U.S. Government Printing Office.

Department of Defense. (2003b). *Transformation Planning Guide*. Washington, DC: U.S. Government Printing Office.

Department of Defense. (2003c). *The Defense Acquisition System*. Department of Defense Directive. Number 5000.1. Washington, DC: U.S.Government Printing Office.

Department of the Navy. (1999). *Verification, Validation, and Accreditation (VV&A) of Models and Simulations*. Secretary of the Navy Instruction 5200.40.

Department of the Navy. (2002a). *Navy Modeling and Simulation (M&S) Management*. Office of the Chief of Naval Operations Instruction 5200.34.

Department of the Navy. (2002b). *Resource Management Process.* Secretary of the Navy Memorandum of November 26, 2002- FY 2005.

Department of the Navy. (2003a). *Accreditation and Use of Performance/Pricing Models in POM-06.* Washington, DC.

Department of the Navy. (2003b). *Aviation Depot Maintenance Model Accreditation: Airframes and Engines.* Washington, DC.

Department of the Navy. (2003c). *Navy Program Development Procedures for Program Review (PR-05).* Memorandum Serial 801C/3U638000.

Department of the Navy. (2003d). *Performance/Pricing Models Verification & Validation (V&V) Template for ADRAMS.*

Department of the Navy. (2003e). *Performance/Pricing Models Verification & Validation (V&V) Template for EDRAMS..*

Department of the Navy. (2004a). *Guidance for the Preparation and Submission of the FY 2006/FY2007 Program/Budget Estimates for the Department of the Navy (DON) Program/Budget Review.*

Department of the Navy. (2004b). *Modeling and Simulation Verification, Validation, and Accreditation Implementation Handbook: Volume 1 VV&A Framework*.

Department of the Navy. (2004b1). *Navy POM-06 Process*. Internal Memorandum Serial N801/ 2004U797200.

Department of the Navy. (2004c*). Performance/ Pricing Models Verification & Validation (V&V) Template (NPTL).*

Department of the Navy. (2004d). *Performance/Pricing Models Verification & Validation (V&V) Template (Restoration Model).*

Department of the Navy. (2004e). *Performance/Pricing Models Verification & Validation (V&V) Template.*

Department of the Navy. (2005). *Guidance for the Preparation and Submission of the FY 2007 Program/Budget Estimates for the Department of the Navy (DON) Program/Budget Review.*

Epstein, M. J., & Birchard, B. (2000). *Counting What Counts: Turning Corporate Accountability to Competitive Advantage.* Cambridge, Massachusetts: Perseus Books.

General Accounting Office. (1998). *Intercity Passenger Rail: Financial Performance of Amtrak's Routes*, (Report RCED-98-151).

General Accounting Office. (1999). *Performance Plans: Selected Approaches for Verification and Validation of Agency Performance Information*, (Report GGD-99-139).

General Accounting Office. (2001). *Results-Oriented Budget Practices in Federal Agencies*, (Report No. 01-1084SP).

General Accounting Office. (2002). *Medicare Financial Management: Further Improvements Needed to Establish Adequate Financial Control and Accountability*, (Report AIMD-00-66).

General Accounting Office. (2003). *Department of Defense: Status of Financial Management Weaknesses and Progress Toward Reform*, (Report No. 03-931T).

Kaplan, R.S., & Norton, D.P. (1996). *The Balanced Scorecard: Translating Strategy into Action.* Boston, MA: Harvard Business School Press.

Laack, D.R., & Muessig, P.R. (1997). Optimizing the Selection of VV&A Activities: A Risk/ Benefit Approach. *Proceedings of the 1997 Winter Simulation Conference*. [CD-ROM] (OmniPress).

Lee, C. (2004). GAO Finds Fault With the Federal Books. *Washingtonpost.Com*. Retrieved March 30, 2005, from http://www.washingtonpost.com/wp-dyn/articles/A14775-2004Dec20.html.

Malina, M. A. & Selto, F. H. (2004). Choice and change of measures in performance measurement models. *Management Accounting Research* 15, 441-469.

McCaffery, J. L. & Jones, L. R. (2001*).* *Budgeting and Financial Management in the Federal Government*. Greenwich, CT: Information Age Publishing.

Mercer, J. (1994). The Performance Management and Budget System of the City of Sunnyvale, California. Retrieved March 30, 2005, from http://www.john-mercer.com/bpi_reform.htm.

Mercer, J. (2001). The Results Act: Has It Met Congressional Expectations? In Statement of John Mercer Before the Subcommittee on Government Efficiency, Financial Management, and Intergovernmental Relations U.S. House of Representatives. Retrieved March 30, 2005, from http://www.john-mercer.com/bpi_reform.htm.

Mercer, J. (2002). Performance Budgeting for Federal Agencies: A Framework. Retrieved March 30, 2005, from http://www.john-mercer.com/bpi_reform.htm.

Mercer, J. (2003). Cascading Performance Budgeting. Retrieved March 30, 2005, from http://www.john-mercer.com/bpi_reform.htm.

Myers, A. G. (2004). *Performance/Pricing Model Way Ahead For CY05 and PR07/POM08 Memorandum*.

Navy Modeling and Simulation Management Office. 2005. Retrieved from http://navmsmo.hq.navy.mil/

Office of Information Technology. (1995). *Activity Based Costing. In Business Process Improvement Reeingineering* (Chap. 5). Retrieved March 30, 2005, from http://www.faa.gov/ait/bpi/handbook/chap5.htm.

Office of Management and Budget. (2002). *The President's Management Agenda*. Washington, DC: U.S. Government Printing Office.

One Hundred Third Congress of the U.S. (1993). *The Government Performance and Results Act of 1993*. Washington, DC: Government Printing Office.

Page, E. H., Canova, B. S., & Tufarolo, J. A. (1997). A Case Study of Verification, Validation, and Accreditation for Advanced Distributed Simulation. *AMC Transactions on Modeling and Computer Simulation, volume 7, no. 3*, pages 393-424.

Personal Interviews. (2005). Office of the Chief of Naval Operations.

PricewaterhouseCoopers. (2001). *CTC Accreditation Recommendation*. Internal document.

Sargent, R. G. (2000). Verification, Validation, and Accreditation of Simulation Models. *Proceedings of the 2000 Winter Simulation Conference.* [CD-ROM] (OmniPress).

Shank, J. K. (1993). *Strategic Cost Management: A New Tool for Competitive Advantage.* New York, N.Y. : The Free Press.

Secretary of the Navy. (1999). Verification, Validation, and Accreditation (VV&A) of Models and Simulations. Washington, DC: U.S. Government Printing Office.

Simmons, R. (2000). *Performance Measurement & Control Systems for Implementing Strategy*. Upper Saddle River, New Jersey: Prentice Hall.

U.S. Code: Title 10, 2466. *Limitations on the Performance of Depot Level Maintenance of Materiel.* Retrieved April 2005 from http://straylight.law.cornell.edu/uscode/search/display.html?terms=depot&url=/uscode/html/uscode10/usc_sec_10_00002466----000-notes.html.

Yin, Robert K. (1994). *Case Study Research: Design and Methods*. Thousand Oaks: Sage Publications.

THIS PAGE INTENTIONALLY LEFT BLANK

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California

3. Vice Admiral (Ret) Thomas Hughes (GB/GH)
   Conrad Chair of Financial Management
   Graduate School of Business and Public Policy
   Naval Postgraduate School
   Monterey, California

4. Commander Philip J. Candreva (GB/PC)
   Graduate School of Business and Public Policy
   Naval Postgraduate School
   Monterey, California

5. Captain (Ret) John E. Mutty (GB/MU)
   Graduate School of Business and Public Policy
   Naval Postgraduate School
   Monterey, California

6. Professor Keebom Kang (GB/KK)
   Graduate of Business and Public Policy
   Naval Postgraduate School
   Monterey, California

7. LCDR Pamela S. Theorgood, SC, USN
   Philadelphia, Pennsylvania